WHEN VARIABILITY MATTERS IN SECOND LANGUAGE WORD LEARNING: TALKER VARIABILITY AND TASK TYPE EFFECTS

by

Asmaa K. Shehata

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

The University of Utah

August 2013

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Asmaa K. Shehata**

has been approved by the following supervisory committee members:

Rachel Hayes-Harb	, Chair	**4/29/2013** Date Approved
Jane Hacking	, Member	**4/29/2013** Date Approved
Johanna Watzinger-Tharp	, Member	**4/29/2013** Date Approved
Bruce Smith	, Member	**4/29/2013** Date Approved
Aaron Kaplan	, Member	**4/29/2013** Date Approved

and by **Edward Rubin** , Chair of

the Department of **Linguistics**

and by Donna M. White, Interim Dean of The Graduate School.

ABSTRACT

This study addressed the role of talker variability in the perception of nonnative contrastive phonemes by adult second language (L2) learners who had no prior knowledge with the target language. Specifically, the study explored how training with varying talkers could affect native English speakers' acquisition of the Arabic pharyngeal-glottal contrast, which is not distinctive in their native language. The present study also examined the effects of task type on learners' word recognition ability.

To accomplish this, the present study included two main experiments: Experiment 1 (nonlexical task) and Experiment 2 (lexical task). Sixty adult native speakers of English (with no Arabic experience) participated in the two experiments, 30 subjects in each experiment who were randomly assigned to either a single- or multiple-talker word learning groups. Subjects in the two experiments were presented with nine nonword minimal pairs where six pairs contrasted the Arabic /ħ/ and /h/ phonemes and three pairs included familiar sound contrasts (i.e., /s/ and /ʃ/). The nine nonword pairs were assigned to pictures indicating their meanings and subjects learned the nine nonword pairs in the training phase and were then tested on them later in the testing phase.

Findings of Experiment 1 demonstrated a significant effect of training type ($p < .001$), a significant effect for item type ($p < .001$), and a significant interaction of training type and item type ($p < .001$) for subjects in the multiple-talker environment. That is, their performance was more accurate (91.5%) than the single-talker group

(67%). The same significant findings were found in Experiment 2 where again, subjects in the multiple-talker training group performed more accurately on test items better than their counterparts in the single-talker training group (single-talker group = 65%; multiple-talker group = 87%).

Overall, the results of this experiment provided evidence that multiple-talker training did have a significant effect on the subjects' recognition of the target contrast in a nonlexical discrimination task with above 88% average accuracy. Findings also provided evidence supporting learners' ability to establish lexical representations for the newly learned words that included the target Arabic contrasting phonemes with above 83% average accuracy for only the multiple-talker training group. Even though subjects' scores differed on the two discrimination tasks, this difference was found to be statistically insignificant. That is, subjects' ability to discriminate the novel contrasts was the same on the lexical task as on the nonlexical task regardless of the two tasks' distinct demands.

Findings of the two experiments imply that variability in talkers can contribute to acquiring nonnative contrasting phonemes. Results are considered in relation to their implications for understanding the learning process of L2 novel phoneme contrasts and their lexical processing.

In loving memory of my dearest Mum and to my gracious Dad

with deep appreciation and love. Thank you for all that you taught me about the world

and all you did for me throughout my whole life. To you two, I not only owe success

but also my life.

TABLE OF CONTENTS

LIST OF TABLES

.

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express my very great appreciation to my PhD advisor, Professor Rachel Hayes-Harb, a mentor whose classes, intellectual guidance, continued support, patience, and inspiration have made me a better researcher and a better scholar. Her expertise and kind support have helped me to grow personally and educationally. I could never have asked for a better mentor. Rachel, I am deeply indebted to you more than you will ever know. I would also like to offer my especial thanks to my committee members, Dr. Jane Hacking, Dr. Johanna Watzinger-Tharp, Dr. Bruce Smith, and Dr. Aaron Kaplan, for all their advice and suggestions.

I would like to extend my appreciation to all faculty who generously supported me through the time I spent at the U, including Dr. MaryAnn Christensen, Dr. Mushira Eid, Dr. Fernando Rubio, Dr. Aniko Csirmaz, and many others who I might have missed mentioning. My special thanks also go to my dear colleagues and close friends who I met at the U for helping me get through the difficult times, and for all the camaraderie, entertainment, emotional support, and caring they provided. I am also grateful to all students who enthusiastically participated in my study.

My deepest and warm gratitude goes to my family members for their continued support, words of encouragement, and constant belief in me since day one, especially my father to whom I am deeply indebted for his unwavering love, patience, understanding, unconditional support, and being the trustworthy and open-minded

person that he is. I am also blessed with the best brother in the world, who is the genuine soul that is always guiding, encouraging, and lighting the path for me. I am also grateful to my cute little niece whose constant smile and unfailing cheerfulness reminded me that every day is a new day and that was an endless source of motivation during the writing process. Above all, I would like to thank my Creator for endowing me with health, patience, and knowledge to complete this work and Whose many blessings have made me who I am today.

CHAPTER 1

INTRODUCTION

Speech produced by second language (L2) learners who acquire their second language after childhood is commonly characterized by a foreign accent (Asher & Garcia, 1969; Flege, MacKay, & Meador, 1999; Leather & James, 1991; MacKay, Flege, & Imai, 2006; MacKay, Meador, & Flege, 2001; Mayo, Florentine, & Buus, 1997; Munro, Flege, & MacKay, 1996). Due to its significance as an acknowledged characteristic of L2 learner speech, the issue of foreign accents draws the attention of several scholars in the field of L2 phonology who have become interested in exploring the reasons behind the persistence of the foreign accent in L2 speech. In this regard, previous second language acquisition (SLA) research findings have shown numerous factors that contribute to the accentedness of L2 speech such as age of arrival in a new host country (Baker & Trofimovich, 2006), amount of exposure to the target language (Bradlow & Bent, 2008; Celce-Murcia, Brinton, & Goodwin 1996; Kenworthy, 1987), learners' cultural attitudes (Moyer, 2007), musical ability (Sleve & Miyake, 2006; Wong, Skoe, Russo, Dees, & Kraus, 2007), and length of residence (Flege & Liu, 2001).

1.1 Overview

Novel L2 speech contrasts, which have no equivalents in learners' native language, have been generally recognized as one of the main factors contributing to the complexity of L2 accented speech (Escudero, Hayes-Harb, & Mitterer, 2008; Flege & MacKay, 2004; Flege, Munro, & Mackay, 1995; Hayes-Harb, 2002; MacKay, Meador, & Flege, 2001). Considerable speech perception research reveals that adult L2 learners experience difficulty in learning novel L2 phoneme contrasts that differ from their first language (Best, McRoberts, & Sithole, 1988; Best & Strange, 1992; Goto, 1971; Hayes-Harb, 2007; MacKain, Best, & Strange, 1981; Werker, Gilbert, Humphrey, & Tees, 1981). For example, Japanese learners of English experience difficulty perceiving the English /r/–/l/ contrast, (e.g., *right* vs. *light*), which does not exist in their native language (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Best, 1995; Boatman, 1990; Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Bradlow & Pisoni, 1999; Goto, 1971; Hattori & Iverson, 2009; Logan, Lively, & Pisoni, 1991; Mochizuki, 1981; Strange, 1995; Takagi, 2002).

Likewise, native Dutch speakers display an inability to differentiate the English /æ/-/ɛ/ contrast, e.g., they have difficulty with the English words *med* and *mad* (Broersma, 2005; Cutler & Broersma, 2005). Similarly, native English speakers have difficulty discriminating contrasts, such as Hindi retroflex /ɖ/ versus dental stops /d̪/ (Werker, Gilbert, Humphery, & Tees, 1981; Werker & Tees, 1984). Prior research, on the other hand, found evidence that lab training may improve L2 learners' perception of novel L2 contrasting phonemes (Bradlow, Pisoni, Yamada, & Tohkura, 1997).

Many studies have shown that adult L2 learners have difficulty establishing distinctive phonetic categories for unfamiliar speech phonemes that are similar to their native language equivalents but are perceived in a different phonetic way (Best, McRoberts, & Sithole, 1988; Flege, 1987). However, it is not yet well-understood why a foreign accent is so persistent in L2 learners' speech. Thus, more studies that further explore this issue can help us build a better understanding of the persistence of accent in L2 speech.

1.2 Proposal and Organization of This Dissertation

The main objective of the current study is two-fold: (1) to examine whether talker variability plays a role in the acquisition of the Arabic pharyngeal-glottal phoneme contrast by native English speakers with no Arabic experience, and (2) to explore the possible effect of task type (i.e., nonlexical versus lexical) on learners' acquisition of novel phoneme contrasts.

The present dissertation includes eight main chapters that are organized as follows. Chapter 2 is an overview of empirical studies of L2 speech perception and word recognition that presents prior research on unfamiliar phoneme contrasts, including those that examined the relative influence of talker variability on L2 phonological acquisition. Furthermore, this chapter introduces the gap in the literature followed by the research questions and hypotheses.

The study design is presented in Chapters 3 and 4, and results are reported in Chapter 5. First, Chapter 3 presents Experiment 1, which investigated whether talker variability may influence learners' distinction of unfamiliar phoneme contrasts on a nonlexical task. Chapter 4 displays Experiment 2 that used a lexical task to explore the role of variation in talker's voice on L2 learners' lexical processing of novel L2 contrasts. Chapter 5 introduces the main results of both Experiment 1 and Experiment 2. It also represents a comparison of findings of the two experiments.

Chapter 6 synthesizes the main findings from Chapter 3 and Chapter 4 and proposes theoretical pedagogical implications of the main findings. Chapter 7 provides a brief summary of this dissertation. Finally, Chapter 8 addresses the study's limitations and proposes directions for future research.

CHAPTER 2

LITERATURE REVIEW

2.1 Background

A large number of empirical studies have addressed the phenomenon of accented speech through focusing on the learning of nonnative phoneme contrasts, which involves learning distinctive features that distinguish them. These studies have tended to investigate this issue in two ways: cross-language research exploring L2 learners' acquisition of a second language at different stages (Curtin, Goad, & Pater, 1998; Cutler, Weber, & Otake, 2006; Escudero, Hayes-Harb, & Mitterer, 2008; Flege, Bohn, & Jang, 1997; Hayes-Harb & Masuda, 2008; Ota, Hartsuiker, & Haywood, 2009; Pállier, Bosch, & Sebastián-Gallés, 1997; Pater, 2003; Pater, 2004; Strange, Polka, & Aguilar, 1989; Weber & Cutler, 2004), and training studies in which L2 learners have been trained in laboratory settings to learn unfamiliar speech contrasts over the course of an experiment session (Bradlow, Pisoni, Yamada, & Tohkura, 1997; Logan, Lively, & Pisoni, 1991; Lively, Logan, & Pisoni, 1993; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994). In this regard, previous research has indicated that training can enhance L2 learners' discrimination of unfamiliar contrasting phonemes that do not have counterparts in their native language.

The following discussion focuses on reviewing prior research to introduce findings of both cross-language and training studies that have been reported in the literature highlighting significant issues with respect to the learning of nonnative contrasting phonetic categories.

2.2 Acquisition of Second Language Lexical Phoneme Contrasts

In an attempt to examine whether adult L2 learners demonstrate an ability to categorize and create lexical representations for the newly learned tokens that included nonnative contrasts, mixed results have been reported in the literature. Some studies have reported the difficulties of phonemic categorization and lexical processing of unfamiliar contrasting phonemes by adult L2 learners (Boatman, 1990; Curtin, Goad, & Pater, 1998; Ota, Hartsuiker, & Haywood, 2009; Pater, 2003; Pater, 2004; Strange, Polka, & Aguilar, 1989) and other studies found that adult L2 learners might not be able to categorize the novel nonnative contrasting phonemes, but they could lexically retain them in their long-term memory (Cutler, Weber, & Otake, 2006; Escudero, Hayes-Harb, & Mitterer, 2008; Weber & Cutler, 2004). The following presents a discussion of these two main findings.

Numerous studies have tested the inability of L2 learners to distinctively categorize L2 contrasting phonemes. For example, Pállier, Bosch, and Sebastián-Gallés (1997) examined whether listeners, who were brought up in a setting where two or more languages were spoken, were able to master the sound systems of these languages and easily move from the native language to the second language or whether they only use one language system to process the other. To explore this question, Pállier, Bosch, and

Sebastián-Gallés chose the Catalan /e/ and /ɛ/ contrast and assessed its perception by 40 participants from Catalonia who were proficient speakers of both the Spanish and Catalan languages. Half of the participants were Spanish-dominant bilinguals, and the other half included Catalan-dominant bilinguals. Three main tasks were used: a classification task, an AX discrimination task, and a typicality judgment task in which subjects heard a word and were asked to pay attention to the initial vowel in the word. Then, subjects listened to a group of vowels that were introduced independently, and they were instructed to rate the degree to which each single vowel matched the initial vowel in the words presented before on a scale from 0 (very bad) to 5 (very good).

Findings showed that both the Catalan and the Spanish groups perceived the target contrast differently. While Catalan-dominant participants perceived the two phonemes as two separate phonemes, Spanish-dominant bilinguals perceived them as the same phoneme. This displayed that Spanish-dominant speakers who were exposed to the Catalan language at an early age were not able to differentiate Catalan words containing /e/ from minimally different words containing /ɛ/ that did not have an equivalent in Spanish. Instead, they used the Spanish /e/ to categorize both members of the distinctive Catalan phonemes. The study concluded that early exposure to a second language was not enough to gain the phonological competence of native Catalan speakers. Although Spanish-dominant speakers in the Pállier, Bosch, and Sebastián-Gallés (1997) study were exposed to the Catalan language at 6 years of age, they did not reach the same phonological competence as Catalan-dominant speakers.

Following up on Pállier et al. (1997), with a very similar population, Pállier, Colomé, and Sebastián-Gallés (2001) conducted another study aimed at exploring lexical processing in the group of Spanish dominant bilinguals who were fluent in Catalan. In this study, Pállier and colleagues recruited 64 undergraduate psychology students from the University of Barcelona. Participants heard words contrasting Catalan /o/ and /ɔ/ and were instructed to indicate as quickly as possible whether each word was a real word or not.

Results indicated that Spanish-dominant bilinguals did display repetition priming for Catalan minimal pairs contrasting phonemes that exist in Catalan but not in Spanish. For instance, the Spanish-dominant bilinguals' response to the Catalan word *dóna* (s/he gives) that included the phoneme /o/ was faster if it was preceded by the word *dona* (woman) that included the phoneme /ɔ/. This finding was explained as evidence that Spanish-dominant bilinguals processed Catalan minimal pairs as homophones of the same phoneme, not as two independent phonemes. The study concluded that Spanish-dominate bilinguals had problems learning the Catalan-specific contrasting phonemes.

Along the same line, Curtin, Goad, and Pater (1998) explored how both monolingual native English and French speakers perceived voicing and aspiration contrasts in Thai. Three groups of subjects participated: 9 monolingual native speakers of Canadian English, 8 monolingual native speakers of Canadian French, and 10 native Thai speakers. The researchers used 18 words contrasting in onset position. Each word was paired with a picture of a noun for presentation to the participants in two tasks: a picture selection task and an ABX nonlexical task in which three auditory stimuli were presented and the participants' task was to choose which of the first two was matched by the final stimulus (X).

It was found that both English and French speakers performed better on contrasts in voicing than in aspiration. Even though native English speakers were able to discriminate the novel Thai voicing and aspiration contrasts on the ABX nonlexical, they were unable to distinguish them in the picture selection task that required sensitivity to the target contrasts.

In a follow-up study to Curtin, Goad, and Pater (1998), Pater (2003) examined the role of task effects on English speakers' performance on both lexical and nonlexical discrimination tasks with respect to voicing and aspiration. Findings exhibited the opposite pattern to the one reported in the study of Curtain et al. (1998), where participants performed more accurately on contrasts of aspiration than voicing in the lexical task. This supported the claims made in earlier voice onset time (VOT) investigations. Furthermore, unlike findings from previous research that demonstrated the different demands of each task type to result in different outcomes, performance of subjects in Pater (2003) was consistent on the two tasks (76% correct on both tasks).

Another important finding of Pater's study was the relationship found between the place of articulation, and both voicing and aspiration. While voicing was found to be better discriminated on alveolars, aspiration was better distinguished on the labials. Pater concluded that L2 learners, who could not reliably differentiate between novel distinctive phonemes, might perceive them similarly and might categorize them into a single phoneme. Thus, they could not lexically discriminate the difference between the target-language contrastive phones.

More recently, Ota, Hartsuiker, and Haywood (2009) investigated the role of the first language (L1) phonology in the lexical representations of unfamiliar speech contrasts. To achieve this, Ota and colleagues gave a visual semantic-relatedness decision task to three groups of participants: 20 native English speakers, 20 native Japanese speakers, and 20 native Arabic speakers. In this task, participants were instructed to judge whether or not there was a semantic relationship between the words of each of the shown pair. For instance, native Japanese learners of English judged pairs, such as *lock-hard* and *rock-key*, to be more semantically related than other control items, such as *sock-hard* and *sock-key*. To illustrate, hearing the word *lock* activated the semantic network of *rock* and the opposite. Ota and his colleagues used two main contrasts: the English /l/-/r/ contrast and the Arabic /p/-/b/ contrast, besides filler stimuli.

Their findings proposed that both Japanese and Arabic speakers encountered difficulty in differentiating the nonnative contrast (/l/-/r/ and /p/-/b/), where native Arabic speakers made more mistakes than Japanese speakers. The finding also indicated that even in the absence of auditory perception, L1 phonology impacted the perception of L2 words and their lexical representations. Findings were compatible with those of previous research (e.g., Pallier, Bosch, & Sebastián-Gallés, 2001) that considered L2 learners' ability to reliably categorize unfamiliar speech phonemes as a prerequisite for lexically processing nonnative phoneme contrasts.

The aforementioned studies have consistently claimed that L2 learners encounter difficulties in categorizing and/or establishing lexical representations for the nonnative distinct speech phonemes. Moreover, they assumed that learners' phonetic perception could determine their lexical storage of the unfamiliar phoneme contrasts. In contrast to these findings, other studies have found cases in which adult L2 learners seem to be able to establish lexical representations for the novel contrasts even if they cannot reliably differentiate them in nonlexical discrimination tasks. For instance, Weber and Cutler (2004) used eye-tracking technology to investigate the mapping of phonetic information to lexical entries in second language.

The study included four eye-tracking experiments using stimuli that included the lax vowel pair /ɛ/–/æ/ and the diphthong pair /aɪ/-/eɪ/. Eighty native Dutch-speaking participants (20 participants in each experiment), who were highly proficient English speakers, were instructed to click on pictures, which matched the words they heard, presented on a computer screen. The results indicated that Dutch speakers fixated longer and more frequently on a picture of a *pencil*, for example, when the target word was *panda*, than on the less confusable distractor (e.g., *beetle*).

Furthermore, Dutch speakers activated the word *pencil* that included the familiar vowel /ɛ/ when they heard the unfamiliar vowel /æ/ in the word *panda*, but not the reverse. This activation was described as being asymmetric. Learners maintained a distinction between English words containing /ɛ/ and /æ / in their lexical representations, even though they could not perceive the contrast in the online auditory word identification task. Two explanations for the dominance of /ɛ/ over /æ/were provided: either /ɛ/ is phonetically closer to the single native Dutch vowel, or the orthographic information influences the construction of lexical representations, because only the letter “e” represents a front central vowel in Dutch.

In a replication of Weber and Cutler’s (2004) study, Cutler, Weber, and Otake (2006) carried out a further eye-tracking study with a different language population: Japanese speakers who listened to the English /r/-/l/ contrast. Native Japanese speakers have problems distinguishing these two phonemes that do not exist in their native language, and they tend to identify /l/ more often than /r/ as /ɾ/. Therefore, both /r/ and /l/ were expected to be perceived as /l/ because /l/ is closer to the Japanese /ɾ/. Here, the two authors investigated the following two main questions: if the asymmetry reflecting distinct lexical representations can be replicated for different listener populations other than the Dutch learners of English who participated in Weber and Cutler's (2004), and how to explain the lexical "dominance" of one of the members of the unfamiliar phoneme contrasts over the other.

To answer these two questions, 24 native speakers of Japanese were recruited, and listened to 20 pairs of English words contrasting /l/ and /r/ in the onset position, such as *lady*/*radio*, *ladder*/*radish*, *loaf*/*rose*, and *legs*/*wreck*. Cutler and colleagues found a similar asymmetric pattern of lexical activation to that found in Weber and Cutler's (2004) study, in which participants activated /l/ when they heard a word started with /r/ but not the reverse. Their finding showed that /l/ was the dominant category for lexical activation. Adult L2 learners were able to establish lexical representations for the unfamiliar contrasting phonemes that they were not able to accurately distinguish in a listening task. Nevertheless, the orthographic hypothesis that was presented earlier to explain the direction of the asymmetry was not supported by their data (i.e., /r/ is not the dominant category but /l/ that is not the corresponding grapheme-phoneme found in their native Japanese language).

Considered together, the findings of Weber and Cutler (2004) and Cutler, Weber, and Otake (2006) provided evidence that adult L2 learners could lexically encode novel L2 contrasting phonemes. This finding contradicted the claim of previous studies, such as Curtin, Goad, and Pater (1998) and Pallier, Bosch and Sebastián-Gallés (2001) that claimed L2 learners' mapping of nonnative distinct phonemes as homophones of the same phoneme.

In an attempt to further investigate the validity of the orthographic hypothesis for the direction of the asymmetry that was introduced by Cutler et al. (2006) and did not sustain for native Japanese learners of English /l/ and /r/, Escudero, Hayes-Harb, and Mitterer (2008) did a follow-up study that examined how Dutch learners of English could establish lexical contrasts for unfamiliar distinctive phonemes (i.e., English /ɛ/ and /æ/) that they were unable to distinguish in online listening tasks.

The study mainly asked whether the availability of the spelled forms of nonwords could affect the phonological content of L2 learners' lexical representations. Hence, 50 Dutch-English bilingual speakers were recruited and divided into two groups: one group was presented with the auditory forms and their pictures, while the other group was presented with the auditory forms with the spelled forms and pictures. All participants learned 20 English nonwords. Results demonstrated that the group of participants who saw only the pictures while listening to the auditory forms was not able to reliably categorize the English /æ/and /ɛ/ and had difficulties differentiating words containing them. However, the other group that saw both the spelled forms and pictures and heard audio forms demonstrated a unique response. That is, they had more glances toward /ɛ/ when /æ/ was presented, but not the reverse. Findings imply that the availability of orthographic forms could help participants establish lexical representations for unfamiliar contrasting phones.

Further support for the claim that L2 learners can retain lexical contrasts is provided by Hayes-Harb and Masuda (2008) who examined native English speakers' perceptual capability to lexically process the Japanese consonant length contrast (e.g., *oto* vs. *otto*) in their memory using two different tasks: listening and production tasks. 3 groups of participants were recruited: 12 monolingual native English speakers, 12 native English speakers who studied Japanese for a year, and 12 native Japanese speakers (as a control group). Participants learned 12 Japanese nonsense stimuli (4 had singleton consonants, 4 had geminate consonants, and 4 were filler words) as names of brands in Japanese. Then, they were tested on their ability to match the words and pictures and produce the Japanese words in a naming task.

The results confirmed that even though learners in the three groups correctly distinguished the target Japanese contrasts in the given task, only two groups (i.e., native Japanese speakers and native English speakers with 1 year of Japanese) were able to accurately produce the target Japanese contrasts. It was suggested that adult L2 learners retained the unfamiliar contrasting phonemes as being strange forms of their native L1 phonemes. According to Hayes-Harb and Masuda (2008), native English speakers were able to establish lexical representations for the Japanese geminate consonant /tt/ as a strange English /t*/ sound.

Based on the findings of the studies discussed above, it can be concluded that previous studies that have examined the issue of acquiring nonnative contrasts included learners from a limited number of first language backgrounds such as Japanese and Dutch. In other words, little research has been conducted that explored limited numbers of speech contrasts, such as the English /ɛ/-/æ/ for Dutch speakers and the English /l/-/r/ for Japanese speakers, for some groups of listeners from restricted native language backgrounds. Therefore, the question now is whether results of prior studies are limited only to these groups of L2 learners of English or if they can be extended to other groups of learners and other target contrasting phonemes. Further research is definitely needed to answer this question.

While the research discussed thus far has considered L2 word recognition, there are a number of word-learning studies that have explored the acquisition of unfamiliar phoneme contrasts by L2 learners in a laboratory setting, i.e., training studies. These studies are reviewed in the next section.

2.3 Training Studies

Reviewing the literature on L2 learning and development shows numerous laboratory-based training studies that have examined the effect of training on improving the abilities of L2 learners' perception and production of the novel distinctive phonemes over time (Aliaga-Garcia, 2007; Aliaga-Garcia & Mora, 2009; Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Henning, 1966; Lively, Logan, & Pisoni, 1993; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Logan, Lively, & Pisoni, 1991; Mueller & Niedzielski, 1968; Tajima, Rothwell, & Munhall, 2002). The primary objective of these studies was to identify the precise conditions under which L2 learners can acquire the phonology of the new language. In these studies, both infants and adult L2 learners were trained in a laboratory setting within experimental sessions on some phonological aspects of a second language.

While early previous training studies did not include multiple sources of variability, which were found to positively affect laboratory training studies (Pisoni, 1971;

Sommers, Nygaard, & Pisoni, 1994; Strange, 1972; Strange & Dittman, 1984), there was another group of prior studies described as high-variability training studies, which have incorporated different sources of variability such as stimuli, talkers, phonetic environments and tasks (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Lively, Logan, & Pisoni, 1993; Lively, Logan, & Pisoni, 1994; Lively, Pisoni, & Logan, 1992; Logan, Lively, & Pisoni, 1991; Pisoni, Lively, & Logan, 1994; Yamada, 1993). The following two subsections briefly discuss these two types of training studies and highlight their main findings.

2.3.1 Early Training Studies

This group of studies refers to the early training studies that have explored the impact of only one or two sources of variability (e.g., talker, stimulus, phonetic environments, and speaking rate) on language perception and spoken word identification (e.g., McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002; Pisoni, 1971; Strange, 1972; Strange & Dittman, 1984). For example, Pisoni (1971) conducted six experiments to investigate how L2 learners perceived few English consonants and vowels where subjects heard synthesized English single phonemes and isolated words produced at different formant levels and their task was to identify each phoneme they heard. It was found that subjects' perception of the target stop consonants was poor, due to the inability of learners to store them in their short-term memory compared with vowels.

Similar findings were demonstrated by Strange (1972), who used synthetic sounds to examine the effectiveness of training in improving native English speakers' discrimination of different VOT of the Thai voiced consonants that were introduced in isolated words. In a training that lacked variability, results showed a minor improvement in subjects' recognition abilities. In general, findings displayed training to be insignificant in helping subjects identify the acoustic differences between the target Thai consonants. Strange concluded that training should include a variation of phonetic contexts in order to influence learners' perception of speech categories.

Another study that confirmed the inadequacy of this type of training studies was Strange and Dittman's (1984), where 8 female native Japanese speakers participated. Subjects attended 14-18 training sessions to identify the English /r/ and /l/ in only word-initial position. Strange and Dittman used three sets of stimuli in their training: a set of real word minimal pairs (e.g., *rock*/*lock*), and two sets of synthetic minimal pairs where feedback was provided for each correct response. Their findings revealed the inability of Japanese listeners to display a noteworthy improvement to distinguish /r/ and /l/ in a generalization task with natural speech tokens involving /r/-/l/ minimal pairs. While subjects were able to transfer knowledge acquired during training to identify the unfamiliar phoneme contrast in nonword stimuli, they were not able to do so with naturally produced real words. Subjects' low performance could be due to the lack of stimulus variability in the given training that mainly focused on using short-term memory by utilizing only low-level information in the available speech signal.

In a recent study that followed the design of the early training studies, McCandliss, Fiez, Protopapas, Conway, and McClelland (2002) explored the identification of the English /r/ and /l/ contrast by native Japanese speakers. McCandliss and colleagues used two training environments: a high-variability training in which multiple talkers produced real and nonwords, and a limited-variability training where sequences of phonemes ranging from /r/ to /l/ were spoken by a single talker. Two groups of participants were recruited to participate in the two training environments. Findings indicated that only the performance of participants in the high variability training was significantly improved after training. The authors concluded that the progress in the performance of L2 learners in this group could be due to the availability of the rich training environment that included different types of stimuli spoken by several talkers, which in turn helped learners improve their perceptual identification of the linguistic content of L2 speech.

The studies discussed above shared the following characteristics. First, they trained language learners by using synthetic speech and reported a small improvement, albeit minor, in participants' ability to generalize their perceptual learning of the target contrasts to real words presented in natural speech. Second, they all used a small number of stimuli, talkers, listeners, and limited phonetic environments (i.e., onset position). Thus, it can be concluded that findings of these studies reveal the inadequacy of this type of perceptual training to help L2 learners differentiate the novel contrasts in generalization tasks.

To sum up, the lack of variability in stimuli, talker, and phonetic environments in such training studies has been recognized as a shortcoming in these studies. In the following training studies, researchers have attempted to improve these weaknesses. More details about the second type of training studies are introduced in the following section.

2.3.2 High-Variability Training Studies

To avoid the limitations of perceptual training in early studies, a series of training studies with Japanese speakers demonstrated that training environments could be enhanced to improve Japanese listeners' perception of the English /l/ and /r/ contrasts that were displayed in English minimal pairs (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Lively, et al. 1993; Lively, Logan, & Pisoni, 1994; Lively, Pisoni, & Logan, 1992; Logan, Lively, & Pisoni, 1991; Pisoni, Lively, & Logan, 1994; Yamada, 1993; Wang, Spence, Jongman, & Sereno, 1999; Wang, Jongman, & Sereno, 2003).

As an example, Logan, Lively, and Pisoni (1991), who were the first to develop the high-variability training paradigm, adapted the same stimuli used by Strange and Dittman (1984) to train 6 adult Japanese speakers to distinguish the English /r/ and /l/ contrast in a variety of word positions (initial singleton, initial cluster, intervocalic, final singleton, and cluster) produced by five native English talkers. After a 3-week training period, participants' perception of the English /r/ and /l/ was significantly improved in two generalization tasks that included novel real English words spoken by new talkers due to the high-variability training environment.

The same finding was confirmed by Lively, Logan, and Pisoni (1993) who used two training environments: multiple-talker training and single-talker training that focused on examining which of the three word positions (e.g. initial singleton, initial cluster, and intervocalic) Japanese learners found the most difficult one. Their findings displayed a significant improvement in subjects' identification accuracy for /r/ and l/ in initial clusters. That is to say, native Japanese speakers in the multiple-talker training were more accurate than subjects in the other group at identifying new words spoken by both familiar and unfamiliar talkers on the generalization task. Their findings displayed the beneficial influence of high-variability perceptual training on improving native Japanese speakers' ability to constantly distinguish the English /l/ and /r/ phonemes contrasts.

In a follow-up study, Lively, Logan, and Pisoni (1994) trained Japanese learners of English in Japan for 3 weeks using the same stimulus set of Lively, Logan, and Pisoni (1993). After 15 training sessions, Japanese speakers' identification of the English /r/ and /l/ contrasts was significantly improved as a result of the high variability training paradigm. Subjects' ability to retain the new contrasts was tested through generalization tests that were administered 3 months later. Findings of these tests displayed no significant decline in subjects' ability to reliably categorize the target contrasting phonemes, confirming the efficiency of this type of training in the acquisition of nonnative contrasts.

In a subsequent study, Yamada (1993) utilized 45 training sessions to examine whether monolingual native Japanese speakers would be able to generalize the training experience to new English words produced by unfamiliar talkers. Results displayed that subjects steadily improved in distinguishing English words contrasting /r and /l/ in the initial position. Conversely, the performance of the monolingual native Japanese speakers was lower than native English speakers' performance.

In the same vein, Yu and Jamieson (1993) used the same stimuli of Logan, Lively, and Pisoni (1991) to train native Korean speakers to distinguish the English /r/ and /l/ phonemes. Their findings also confirmed that learners' performance could improve with training. The authors also reported the final position of the target consonant contrasts to be the most challenging phonetic environment for L2 learners. Contradicting researchers' expectation that the initial position of the contrast would be the most difficult, their findings revealed that the target phoneme contrasts occurring in the final position were the most difficult word position for native Korean speakers who encountered difficulties in categorizing the two contrastive English /r/ and /l/ phonemes dependably.

Similar findings were reported by Bradlow, Pisoni, Akahane-Yamada, and Tohkura (1997), whose study reported that perceptual training helped native Japanese speakers improve their production of the English /r/ and /l/. Recently, Wade, Jongman, and Sereno (2007) confirmed the effectiveness of training with variable stimuli, a number of different talkers, and different contextual environments, in improving native English listeners' perception of English vowels in Spanish-accented productions.

Whereas previous studies explored the impact of lab training on L2 learners' perception focusing on unfamiliar consonants within a single binary contrast, Lambacher, Martens, Kakehi, Marasinghe, and Molholt (2005) tested the influence of lab training on both the perception and production of five American mid- and low vowels /æ/, /ɑ/, /ʌ/, /ɔ/, /ɝ/ by two groups of adult native speakers of Japanese who were trained for 6 weeks. Recordings of the two groups' pre-/posttraining of vowel productions of the five vowels were made, and then evaluated by American native English listeners using three tasks: a listening task, an acoustic analysis of the vowel productions, and a production task. The findings confirmed a remarkable development in the recognition performance of the trained native Japanese speakers resulting from the 6-week identification training that was accompanied with feedback. Participants' production of the target American vowels was also improved.

All the studies discussed above have investigated how perceptual training could influence the perception of unfamiliar segmentals. Recent empirical studies, on the other hand, have explored the relative influence of training on the learning of nonnative suprasegmental contrasts such as Mandarin tones. For instance, in a study by Wang, Spence, Jongman, and Sereno (1999), 8 native English speakers were trained for 2 weeks to discriminate the four Mandarin tones in real words spoken by native Mandarin speakers. Wang and colleagues used a pretest and a posttest to check the possible improvement in subjects' identification of Mandarin tones due to training. There were also two generalization tests to investigate whether the training benefit can be extended to new stimuli produced by new talkers. To check the relative influence of long-term training, a long-term retention test was conducted 6 months after the training. Results indicated development in subjects' performance from pretest (69% correct responses) to the posttest (90% correct responses) with 21% increase in subjects' tone detection accuracy.

Wang and colleagues also found that subjects' recognition of Mandarin tones was enhanced in the two generalization tests, where trainees were successfully able to extend their knowledge of the target tone contrasts to new stimuli from novel talkers. The authors concluded that using high-variability training paradigm could improve native English speakers' perception of Mandarin tones. Findings of Wage et al. (1999) consequently demonstrated that training procedures, which have been proved to be successful in improving adult L2 learners' acquisition of novel segmental contrasts, could be beneficial in developing adult L2 learners' perception of the suprasegmental contrasts as well.

Moreover, Wong and Perrachione (2007) explored whether phonological awareness influences L2 learners' acquisition of nonnative phonetic features. Specifically, they asked if nontone language speakers with prior musical experience were more likely to learn to use suprasegmentals (i.e., Mandarin tones) for word identification. In their study, 17 adult native English speakers who had no exposure to a tone language were trained to distinguish three mandarin tones (i.e., 1 (level), 2 (rising), and 4 (falling)) in a lexical task. Researchers excluded the Mandarin Tone 3, which results from previous research found to be the most difficult one for both learners and native speakers of Mandarin. Before training, all subjects participated in a tone awareness test in which they listened to five Mandarin vowels (i.e., /a/, /e/, /i/, /o/, and /y/) produced by four native Mandarin speakers with three resynthesized different Mandarin tones (Tone 1, tone 2, and tone 4) and their task was to recognize the pitch pattern of each stimulus they heard.

In a second task, Wong and Perrachione trained the same group of subjects to identify 18 English nonsense words, which were associated to drawings depicting their meanings. Researchers then had subjects undergo a testing phase in which items tested subjects' ability to distinguish the tone patterns associated with each image. Subjects chose the corresponding image to each of the words they heard. Results showed an increase in subjects' identification accuracy by average of 83.23% where 9 subjects were found to be more accurate than the other 8 subjects who were classified by the authors as 'less successful learners'. Wong and Perrachione concluded that learning suprasegmentals can facilitate the identification of newly learned words that was found to be influenced by learners' prior musical experience.

Using the same methodology, Lee, Perrachione, Dees, and Wong (2007) investigated the influence of stimulus variability on L2 learners' acquisition of phonetic contrasts (suprasegmentals). Researchers trained 47 adult native English speakers in two training conditions: multitalker training and single-talker training, to use Mandarin tones to detect 18 English nonwords in a lexical task. Lee and colleagues found a significant difference between learners' performance in the two training conditions. Subjects in the multitalker training who got high scores in the pitch-identification task improved in their identification of the tone contrasts, whereas subjects with low pitch-identification scores performed more accurately in the single-talker training. Researchers concluded that the use of high-variability training when learning novel words was beneficial to L2 learners only when they are able to store the phonetic details of the given stimuli in memory. This finding suggests a correlation between learners' high sensitivity to the pitch pattern and the positive influence of talker variability in distinguishing novel suprasegmentals contrasts.

In summary, laboratory-based training studies demonstrated that L2 learners' perception and production of nonnative phoneme contrasts could be improved as a result of training (e.g., Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Lively, Logan, & Pisoni, 1993; Lively, Logan, & Pisoni, 1994; Logan, Lively, & Pisoni, 1991). In addition, previous studies highlighted the importance of the high-variability training paradigm, which included variable stimuli produced by multiple talkers in different phonetic environments that allowed a successful perceptual learning. They also indicated that learners' ability to transfer what they learned to new stimuli spoken by new talkers could happen. Based on the robust outcomes of these word-learning studies, the present study will follow the same methodology, but use an unfamiliar L2 phoneme contrast.

In addition to unfamiliar phoneme contrasts and how they are discussed in cross-language studies that have been addressed in section (2.2), and training studies and their two main types, which are introduced in section (2.3), the other main factor investigated in the present study is talker variability. The latter is addressed in more detail in the next section.

2.4 Talker Variability

Differences among talkers create variation in speech signals that is known as one of the principal sources of variability influencing learners' perception (Halle, 1985; Summerfield & Haggard, 1973). There are a number of different elements that could cause differences in qualities of talkers' voices. These include the shape, size, and length of the vocal tract; and how talkers use different acoustic measures, such as rate and length of formant transitions. More important, these elements have been found to be influential in listeners' perception of the spoken language (Carrel, 1984; Fant, 1973; Ladefoged, 1980; Peterson & Barney, 1952).

Earlier research has paid attention to the significance of talker variability, and a review of the literature uncovers a number of studies that have addressed the role that talker variability plays in speech perception in general, and word recognition in particular (Clopper, & Pisoni, 2004; Creelman, 1957; DeCasper & Fifer, 1980; Flege, Mackay, & Meador, 1999; Flege & Fletcher, 1992; Hardison, 2003; Johnson & Mullennix, 1997; Jusczyk, Pisoni, & Mullennix, 1992a; Martin, Mullennix, Pisoni, & Sommers, 1989; Mullennix & Pisoni, 1990; Mullennix, Pisoni, & Martin, 1989; Rost & McMurray, 2009; Sommers, Kirk, & Pisoni, 1997; Sommers, Nygaard, & Pisoni, 1994). Talker variability has been studied in both infant studies (e.g., Barker & Newman, 2004; Houston, Jusczyk, & Tager, 1998; Jusczyk, Pisoni, & Mullennix, 1992; Kuhl, 1979, 1983) and adult studies (e.g., Nygaard, Sommers, & Pisoni, 1994; Nygaard & Pisoni, 1998; Palmeri, Goldinger, & Pisoni, 1993; Yonan & Sommers, 2000). However, these studies reported contradictory findings that are briefly presented in the following subsections that provide a further understanding of how talker variability effects are reported in the literature.

2.4.1 Infant Studies

A thorough review of the literature displays a number of studies that focused on examining the impact of talker variability on infants' perception of novel phonemes. Perhaps the most striking finding was that young infants were found to be more accurate at discriminating unfamiliar phonetic categories produced by different talkers (DeCasper & Fifer, 1980; Hollich, Jusczyk, & Brent, 2002; Houston, 1999; Houston & Jusczyk, 2003; Jusczyk, Pisoni, & Mullennix, 1992b; Rost & McMurray; 2010; Singh, 2008). For example, 6-month-old-infants who first learned fricative contrasts and two other vowel contrasts (e.g., English vowels /a/-/i/, and /a/-/ɔ/), demonstrated some abilities to differentiate the target contrasts when they were spoken by different speakers (Barker & Newman, 2004; Kuhl, 1979, 1983).

More compelling evidence for the positive role of multitalker variability is provided by another group of studies. For example, Rost and McMurray (2009) used a different methodology from prior studies to investigate infants' discrimination of phoneme contrasts; they used pictures whose labels were read by either one talker or a number of different talkers. Their findings demonstrated that infants who listened to labels (e.g., *buk* versus *puk*) spoken by multiple talkers were more accurate at discriminating the difference between words contrasting /p/ and /b/ in initial position than infants who listened to labels spoken by a single talker. Rost and McMurray's (2009) findings demonstrated the effectiveness of using talker variability in discriminating novel phonetic contrasts by infants.

In a similar study, Houston and Jusczyk (2000) examined the effects of talker variability on recognition of words in fluent speech by two different age groups of infants: 7-and-a-half-month-old and 10-and-a-half-month-old infants. When there was 1-day delay between the training and testing sessions, 7-and-a-half-month-old infants demonstrated a significant improvement in their word recognition ability only when stimuli were spoken by talkers of the same gender as the talker in the training session. While these infants were successfully able to generalize training to two female talkers, they were unable to recognize words produced by two new male talkers. This finding suggests that listening to several talkers of different genders did impact the perceptual identification of the spoken words by infants at this age. Nevertheless, 10-and-a-half-month-old infants performed differently as they were able to generalize words produced by a single talker to other talkers of the opposite gender.

In some cases, on the other hand, infant studies showed that talker variability hinders infants' speech recognition performance. In a subsequent study, Jusczyk, Pisoni, and Mullennix (1992b) found that variability in both tokens and talkers interfered with the word recognition abilities of 2-month-old infants after a delay interval. Thus, infants in both single and multiple-talker groups were able to notice the change in the phoneme from *bug* in the training session into *dug* in the test session in the first experiment. Conversely, when the subsequent experiments included a 2-minute delay between training and testing, only young infants in the single-talker group were able to observe the change in the target phoneme. Findings displayed that 2-month-olds were able to create lexical representations for the target contrasting phonemes more robustly when they heard them produced by a single-talker group.

Similarly, Schmale and Seidl (2009) examined in six experiments how variability in voice and accent could affect word recognition abilities of two different age groups of infants: 9-month-old infants and 13-month-old infants. Findings displayed the inability of 9-month-old-infants to detect words when produced by different talkers who varied in both voice and accent. Conversely, 13-month-old infants were able to transfer training to tests that included English words spoken by both novel native English and Spanish-accented talkers. Findings showed that talker variability impeded 9-month-olds' detection of spoken words since it raised the word processing load that in turn resulted in infants' low word recognition performance.

Taken together, it can be concluded that prior research has reported paradoxical findings concerning the role of talker variability in infants' speech perception and word recognition ability. The following section, however, illustrates how variability in talkers affects adults' language accusation, and how this issue has been displayed in the literature.

2.4.2 Adult Studies

A considerable amount of recent research indicates that, like infant research, adult studies have demonstrated a set of contradictory results regarding talker variability and its impact on learners' identification of the speech content. The majority of studies on talker variability have been conducted on adult L1 speakers. However, L2 research shows a very small number of studies that examined the role of talker variability in L2 learners' perception of L2 contrasts. Findings from both L1 and L2 adult research are presented in the following sections.

2.4.2.1 Adult L1 Studies

A number of studies have reported that talker variability reduces learners' performance in listeners' word recognition, naming, and recall paradigms (Church & Schacter, 1994; Creelman, 1957; Flege & Fletcher, 1992; Flege, Mackay, & Meador, 1999; Goldinger, 1992; Goldinger & Pisoni, 1991; Goldinger, Pisoni, & Logan, 1991; Martin, Mullennix, Pisoni, & Sommers, 1989; Mullennix & Pisoni, 1990; Mullennix, Pisoni, & Martin, 1989; Nusbaum & Morin, 1989; Nusbaum & Morin, 1992; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994; Nygaard, Sommers, & Pisoni, 1995; Schacter & Church, 1992; Sheffert & Fowler, 1995; Sommers, Kirk, & Pisoni, 1997; Sommers, Nygaard, & Pisoni, 1994). In one study, Creelman (1957) introduced lists of monosyllabic words produced by 16 talkers in noise to 5 listeners. Findings revealed that listeners were less accurate at identifying words spoken by different talkers than words that were only spoken by a single talker. The authors concluded by emphasizing that talker variability played a negative role in distinguishing the spoken word.

Replicating Creelman' s study, Mullennix, Pisoni, and Martin (1989) investigated how talker variability could be influential in learners' perceptual performance using word lists. To test this claim, Mullennix and colleagues conducted a series of experiments. Experiment 1 included an identification task where 22 native English speakers were instructed to listen to 68 English words produced in noise by either a single-talker or multiple-talkers. Experiment 2 included a naming task in which 12 native English speakers were asked to name each of the target words produced by multiple-talkers once they heard it.

In Experiment 3, 70 native speakers, who did not participate in the previous 2 experiments, listened to 96 English words on a naming task that varied in its word frequency- 48 low frequent words and 48 high frequent words- and produced in 2 environments: a single-talker training and a multiple-talker training. In Experiment 4, however, 30 native English speakers listened to the same stimuli as in Experiment 3, and were asked to write down the words they heard. The general findings of the four experiments revealed that when participants listened to words spoken by one talker, they better identified them than when those words were produced by different talkers - whether these words were high- or low-frequency ones. In other words, subjects performed worse when they listened to different talkers, compared with a single talker, which resulted in a significant improvement in subjects' word recognition ability.

In a similar vein, Martin, Mullennix, Pisoni, and Summers (1989) conducted three experiments including five lists of 10 monosyllabic English words produced by either multiple talkers (high-variability training environment) or by a single talker (low-variability training environment). Findings revealed that listeners' recall accuracy was significantly higher when words were spoken by a single talker as opposed to words produced by different talkers. It was explained that in order for learners to process words spoken by different talkers, they needed additional resources in the long-term memory in comparison with tokens introduced in one talker's voice. This finding supported the insignificant influence of talker variability on L2 learners' perception as it increased the perceptual processing load of the novel tokens that accordingly impeded their acquisition.

In a follow-up study, Mullennix and Pisoni (1990) provided additional evidence for the negative role of talker variability. Researchers used a speeded classification task in which isolated monosyllabic English words were introduced to a group of native English speakers who were instructed to concentrate on words they heard (whether initiated with /b/ or /p/) or talkers' voice (whether it was a voice of a female or a male). Mullennix, and Pisoni's findings highlighted two interesting points. First of all, the authors found that subjects processed words in relation to talkers' voices, suggesting subjects' inability to pay attention to the target phonemes (linguistic information) without also attending to the distinct characteristics (e.g., speech rate) of each talker's voice. This implies the integrity of the linguistic and voice information that cannot be processed independently. Secondly, it was found that those subjects who listened to words spoken by a single talker performed more accurately on the word identification task than the other group who had the target stimuli spoken by different talkers. The findings suggest that familiarity with a talker's voice enhances listeners' discrimination performance.

A further example is Sommers, Nygaard, and Pisoni's (1994) study that tested how talker availability, speaking rate, and amplitude might affect learners' perception performance and whether these three elements would help learners remember novel words. To examine the impact of talker variability, the authors asked two groups of native English speakers in two different training settings: single-talker training and multiple-talker training, to type the words they heard. Likewise, researchers tested two other groups of subjects in either a single speaking-rate (i.e., subjects heard words produced at either fast, medium, or slow rate) or a mixed speaking-rate (i.e., subjects heard words produced at one of the three different speaking rates). Furthermore, Sommers and colleagues recruited 60 more subjects to investigate subjects' performance when they heard word lists differing in either talker variability or speaking rate, and when these word lists differed along these two dimensions. All subjects in Experiment 1 were instructed to type the words they heard as well as they could. All tokens were shown in noise to subjects in different groups.

Findings demonstrated that when a single talker first introduced word lists to learners, they were displayed to accurately identify the target words, as compared to the other listeners who heard the target tokens spoken by different talkers. However, hearing the target stimuli produced by more than one talker at different speaking rates (e.g., low, middle and high) hindered subjects' perception due to too much variability in the given speech signals of each talker.

In short, the results of the aforementioned studies demonstrate that talker variability negatively affects learners' perception of isolated phonemes or words, which is interpreted in terms of the complexity of the input provided in multiple talker environments where the acoustic characteristics of talkers' voices require a considerable mental effort that is found to be a barrier impeding learners' identification. On the contrary, learners show evidence of accurately perceiving the target stimuli when they are introduced by one single talker.

2.4.2.2 Adult L2 Studies

In contrast to the studies discussed above, L2 studies have shown that talker variability, in comparison with single-talker environments, plays a positive role in learning novel L2 phoneme categories. In their study discussed earlier, for example, Logan, Lively, and Pisoni (1991) showed that in contrast to single-talker training, training that included multiple talkers resulted in improving Japanese speakers' acquisition of the English /l/-/r/ phonemic contrasts. The authors confirmed that the availability of different talkers did help learners retain the phonetic information of the target contrast, as they stated; "nonnative listeners encode detailed talker-specific information and apparently store this information in long-term memory" (p. 881). Their findings provide evidence for the retention of talker information in learners' long-term memory that subsequently enabled them to detect new words in the generalization test.

In a similar investigation, Lively, Logan, and Pisoni (1993) demonstrated training with multiple-talkers to be more effective than training with a single-talker. This finding showed the positive effect of talker variability on native Japanese listeners' perception of the English /l/ and /r/ phoneme contrasts. The authors concluded that hearing a single talker did not enable listeners to generalize their word familiarity to tokens produced by new talkers compared with performance of subjects in the multiple-talker training environment.

Further evidence in accord with this view is provided by Wang, Spence, Jongman, and Sereno (1999), who trained native English speakers on Mandarin distinctive tones using real words from different talkers. Their findings showed a statistically significant improvement in listeners' discrimination ability on the target tones. That is, after 2 weeks of trainings, listeners were able to generalize training to new words produced by new talkers. Based on their findings, the authors concluded that the rich input provided in the high variability training was the reason behind subjects' accurate performance in the generalization test.

Recently, Bradlow and Bent (2008) examined the influence of talker variability on learners' transcription skills. Authors gave English sentences produced by three groups of native Chinese talkers: multiple-talker, single-talker, and control environments. Findings revealed the capability of learners in the multiple-talker group to transcribe the target sentences more accurately than the other two groups. Researchers concluded that the beneficial role of talker variability could be extended to accented-speech where it played an advantageous part in improving learners' transcription skills.

Unlike the studies discussed above, Hardison (2003) found the influence of talker variability to be of a marginal significance, where both multiple and single-talker trainings were found to facilitate comprehension of unfamiliar speech. Hardison's study aimed at exploring the impact of word position, adjacent vowel, talker variability, and training type (auditory versus visual) on native Japanese and Korean speakers' perception of the English /r/ and /l/ contrasts. The researcher recruited 16 native Japanese speakers and 8 native Korean speakers who participated in two different experiments. Experiment 1 included two training environments: the first one included auditory and visual inputs and the second training environment only included an auditory input.

Experiment 1 included the following main sessions: pretest, tainting, posttest, and two generalization tests (one with a familiar talker from the training phase and one with an unfamiliar talker). Experiment 2 examined the impact of visual input on training Korean learners of English. Like Experiment 1, the second experiment included two training environments: visual and auditory training group and the auditory only training group. Moreover, each of these training groups was divided into two groups: one where subjects listened to stimuli presented by multiple talkers and another one where subjects listened to stimuli spoken by a single talker. While findings indicated a significant impact of training type, word position, and adjacent vowel on the perception and production of /r/ and /l/ by the 2 ESL participants, they also revealed a marginal significance of talker variability on subjects' performance in the two generalization tests. Korean speakers displayed a minor success in generalizing the training they received to new tokens produced by unfamiliar talkers.

In summary, two essential conclusions can be drawn from the investigations reviewed above. First, testing talker variability and its relative effects has demonstrated mixed findings. While some studies have found talker variability to be an ineffective factor that impairs learners' performance, other studies have provided evidence in favor of the positive role of talker variability. A third group of prior studies, on the other hand, displayed a minor role for talker variability where both single- and multiple-talker training could facilitate learners' comprehension of L2 speech phonemes. These conflicting findings clearly demonstrate the need for conducting further research to address this significant issue.

Second, previous L2 research that has explored the influence of talker variability as related to novel L2 phoneme contrasts mainly used nonlexical tasks that examined learners' online perception of the newly learned contrasts and paid little attention to the lexical processing of these contrasts. Thus, in order to build on findings of previous studies and provide a better understanding of this issue, the main objective of the current study is to further examine the relative impact of talker variability on adult L2 learners' ability to categorically discriminate and lexically store unfamiliar speech contrasting phonemes using both lexical and nonlexical tasks.

2.5 The Significance of the Mental Lexicon and Lexical Representations in Language Acquisition

The term "mental lexicon" that was first introduced by Oldfield (1966) is known as an "intrinsic device" that changes the sound input into meaning. Lexical access, on the other hand, is defined as "the process of formulating an appropriate input and mapping it into an entry in the lexicon's store of sound images matched with their meaning" (Cutler, 1989, p. 342). In an attempt to explain the lexical processing, several models of speech perception have been proposed, such as LAFS (Lexical Access From Spectra) Model (Klatt, 1979) and Trace Model (Elman & McClelland, 1984). According to these models, speech acoustic signals are stored in the learner's lexicon in the form of lexical representations, a middle stage between phoneme categories and meanings that includes all information about the target words, such as spelling, pronunciation, morphological structure, meaning, and syntactic category. Creating lexical representations of the given speech, therefore, helps listeners in interpreting and in turn contributes to their language comprehension.

Due to its significance, the acquisition of lexical phonological representations has become a key topic in the domain of L2 phonology. Unlike cross-language studies that have examined the lexical storage of novel L2 contrasts in several prior studies (Curtin, Goad, & Pater, 1998; Cutler, Weber, & Otake, 2006; Escudero et al., 2008; Hayes-Harb & Masuda, 2008; Pater, 2003; Salverda, Dahan, & McQueen, 2003; Shatman & McQueen, 2006; Weber & Cutler, 2004), training studies have paid comparatively little attention to this issue that was mainly investigated in infant training studies examining the lexical representations of speech phonemes for infants during and after their first year using lexical tasks such as lexical decision, naming-latency, and serial or continuous lexical decision (Lucas, 2000). For example, Jusczyk and his colleagues (Bertoncini, Bijeljac, Jusczyk, Kennedy, & Mehler, 1988; Jusczyk, Bertoncini, Bijeljac-Babic, Kennedy, & Mehler, 1990; Jusczyk & Derrah, 1987; Jusczyk, Jusczyk, Kennedy, Schomberg, & Koenig, 1995; Jusczyk, Kennedy, & Jusczyk, 1995) provided evidence that infants could successfully discriminate familiar syllables by the age of 2 months.

On the other hand, older infants were reported to perform better on segmentation and encoding of familiar and unfamiliar syllables. These findings, moreover, demonstrated the ability of young infants to develop detailed syllabic representations of speech. Additionally, 2-month-old infants displayed an ability to establish lexical representation in their memory for both unfamiliar consonants and familiar vowels in a set of consonant-vowel (CV) sequences over a short period of time (Jusczyk, 1990), and remember a series of three CV syllables that included unfamiliar consonants and vowels (Jusczyk et al., 1995). Furthermore, Jusczyk, Jusczyk, Kennedy, Schomberg, and Koenig (1995) provided evidence that 2- and 3-month-old infants could lexically process bisyllables when they were presented in CV sequences.

Yet the nature of lexical representations and how they are established have not received the attention commensurate with their essential role in adult L2 training studies in general, and high-variability training studies as related to nonnative contrasting phonemes in particular. Instead, prior adult training studies have mainly measured the perceptual performance of L2 learners using nonlexical tasks, such as identification and discrimination tasks (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Lively, Logan, & Pisoni, 1993; Lively, Logan, & Pisoni, 1994; Logan, Lively, & Pisoni, 1991). Therefore, the present study attempts to fill in this gap in the literature by using a lexical discrimination task in the investigation of the role of talker variability in the lexical processing of novel L2 contrasting phonemes.

There are at least two methods for conducting a lexical XAB discrimination task that have been reported in the literature. In the first method that uses sound-picture-picture (SPP), A and B are both pictures presented to subjects whose is task to match the auditory stimulus (X) to the correct picture. In the second method that includes picture-sound-sound (PSS), subjects are presented with two auditory stimuli (A and B) and one visual stimulus (X), and are instructed to match the visual stimulus (X) to the correct auditory one. Pater (2003) provided evidence that these two types of lexical XAB tasks can elicit different results. While subjects performed well on the XAB lexical task where A and B were auditory words, they performed poorly and their abilities to match the word to the correct picture was at chance level on the XAB lexical task where A and B were pictures (Pater, 2003).

This finding can be interpreted in light of Hayes-Harb and Masuda's (2008) possible explanation for how novel L2 contrasts are initially stored in L2 learners' lexicon. According to these researchers, when native English speakers were first introduced to the novel L2 contrast, they probably store the novel phoneme as a 'strange' form of the nearest native language phoneme. Therefore, it was possible that when subjects in Pater's (2003) study were presented with the two contrasting words in the PSS task, they did not store the given contrast correctly. Instead, they were able to compare the two auditory forms in that task and decide which of them was different from the closest L1 phoneme.

However, when subjects were presented with two pictures and one auditory word in the SPP task, they were not able to determine which of the two pictures was associated with the word they heard since the novel contrast was not stored. Consequently, their performance was at the chance level. It can be concluded that the PSS lexical task probably tested learners' discrimination of the difference between the two contrasting phonemes, but did not require the storage of the novel contrast that was definitely required by the SPP lexical task. For this reason and for comparison purposes, the XAB lexical discrimination task (sound-picture-picture) that has been used by Pater (2003) is selected to be used in the present study.

2.6 Effects of Task Type on Learners' Acquisition of Novel Phoneme Contrasts

It has been documented in the domain of L2 phonology that learners' perceptual performance differs according to the type of tasks they are exposed to during training and testing phases (Logan & Pruitt 1995; Matthews & Brown, 2004; Werker & Tees, 1984b). This is due to the different demands of each task type. For example, lexical tasks are found to be more demanding than nonlexical tasks since they require listeners to access their memory for the meaning of the target stimuli (Curtin, Goad, & Pater, 1998).

In their investigation of L2 learners' ability to distinguish and establish lexical representations of nonnative contrasting phonemes, prior SLA studies have used both nonlexical and lexical tasks that have different demands (Curtin, Goad, & Pater, 1998; Hayes-Harb & Masuda, 2008; Pater, 2003). Accordingly, L2 learners' performance differs with respect to L2 phoneme contrasts. For instance, native speakers of English were more accurate at discriminating the Thai voice contrasts after training in a nonlexical task than in the lexical task that required memory of the target contrasts (Curtin, Goad, & Pater, 1998). Conversely, Hayes-Harb, and Masuda (2008) found native English speakers who studied Japanese for 1 year were able to accurately discriminate the Japanese length consonant contrast in the lexical task; however, their performance was significantly less accurate on the nonlexical one.

Unlike the two above-mentioned patterns of previous research findings, Pater (2003) displayed no difference in the performance of native English speakers on both nonlexical and lexical tasks. When subjects were asked to match one of the words they heard to the corresponding picture, as explained earlier in section 2.1, they performed as well on the lexical XAB identification task as they did on the nonlexical XAB identification task. The researcher concluded that the similar design of the two tasks, which included the same pictures and number and types of phases, was behind learners' similar performance on the two different tasks.

In sum, the conflicting findings reported in the literature display a need for further investigation of this issue, which is the second main objective of the current study that examines the possible influence of task type (i.e., nonlexical versus lexical) on learners' recognition of novel L2 contrasts. After clarifying the two main objectives of the present study, the following section explains the reasons for the selected target contrasting phonemes that are used in the present study.

2.7 English Speakers and the Acquisition of the Arabic /ħ/-/h/ Distinction

Reviewing the literature uncovers only a very small number of recent studies that have investigated the acquisition of Arabic language by adult L2 learners in general, and by native English speakers in particular. These few studies have indicated the difficulty of learning Arabic contrasts by native English speakers (Shehata, 2007; Zaba, Bolewicz, & Hayes-Harb, submitted). For instance, Zaba et al. (submitted) examined the perception of pharyngealized and nonpharyngealized Arabic contrasts (i.e., /t/-/$t^ʕ$/ and /d/-/$d^ʕ$/) by native speakers of English when they were presented in three distinctive vowel contexts (i.e., /a/, /i/, and /u/). Researchers used two main tasks: a vowel identification task and an XAB identification task. Their findings showed significant main effects of vowels (p<. 001) and consonants (p=. 015). That is, subjects were able to discriminate the Arabic novel contrasts relying on their familiarity with the English vowel contrasts that helped subjects recognize the target consonant contrasts. In other words, this finding provides evidence for the difficulty of categorizing novel Arabic contrastive phonemes by native English speakers.

A further example is Shehata's (2007) study that investigated how Arabic consonants are perceived by native English speakers. Shehata asked how both learners and teachers of Arabic rate the level of difficulty and/ or easiness of the acquisition of Arabic consonant contrasts at both perception and production levels. It was hypothesized that Arabic learners would encounter difficulties in perceiving and producing Arabic consonant contrasts. Similarly, it was also hypothesized that experience with the target language would play a role in the acquisition of Arabic consonant contrasts. That is, subjects who spent a longer time studying Arabic were expected to encounter less difficulties perceiving and producing its contrastive consonants. On the other hand, learners with short study time were expected to encounter more difficulties.

Two different tasks were included; one was a rating of intelligibility task, and the other a questionnaire. In the first task, 20 native English speakers with different Arabic proficiency levels were recorded producing Arabic consonant contrasts. The intelligibility of each speaker was later rated on a 5-point Likert scale ranging from 1 (=completely unintelligible) to 5 (=completely intelligible) by two groups of listeners: 30 native Arabic speakers and 30 English learners of Arabic.

In the questionnaire, on the other hand, the researcher instructed two groups of listeners (50 native Arabic speakers who are teachers of Arabic and 30 English learners of Arabic) to rate the level of difficulty/easiness of learning different Arabic consonant phoneme contrasts with English equivalents (e.g., /s/-/ʃ/, /k/-/g/, /m/-/n/, and /θ/-/ð/), and with no English equivalents (e.g., /t/-/tˤ/, /d/-/dˤ/, /s/-/sˤ/, /ħ/-/h/, /k/-/q/, /x/-/ɣ/, and /ʔˤ/-/ʔ/) on a 5-point Likert scale ranging from 1 (=completely difficult) to 5 (=completely easy).

Shehata's results (2007) support the hypothesis that Arabic consonant contrasts are difficult to perceive and produce by native English speakers even after years of learning the language. The findings also showed the pharyngeal consonants to be the most unintelligible and difficult phonemes for learners to produce. It was also found that subjects were more accurate at distinguishing the Arabic contrastive consonants that have English equivalents (i.e., /t/ and /d/) than those with no English equivalents (i.e., /k/-/q/). In addition, some pharyngeal consonant contrasts were reported by both teachers and learners to be to be more difficult to discriminate (i.e., /t/-/tˤ/and /ħ/-/h/) than others (i.e., /s/-/ sˤ/ and /d/ and /dˤ/).

Based on these findings, this study chose to pursue the Arabic pharyngeal-glottal contrast for several reasons. First, although there is a rapid increase in the number of learners of Arabic in the United States (US) in which Arabic has been classified as one of the critical languages by the US government, Arabic still remains among the languages that are less studied as a second language (National Security Language Initiative; http://exchanges.state.gov/NSLI)[1].

[1] The link was adopted from Hayes-Harb's grant proposal (2008)

Second, the phonological inventories of Arabic and English have some similarities, but Arabic has a wider range of consonants, in terms of the range of places of articulation, than English does. One key element appearing in Arabic that is missing in English is the phonemic contrast between /ħ/ and /h/. While Arabic includes a pharyngeal place of articulation that occurs in both voiced /ʕ/ and voiceless fricative /ħ/, it also has a voiceless glottal fricative /h/. English, on the other hand, does not have any phones produced at the pharyngeal place of articulation. Instead, English has two phones produced with the glottis (i.e., the voiceless-glottal fricative /h/ and the voiceless-glottal stop /ʔ/).

Third, /ħ/ and /h/ phonemes are positionally unconstrained in Arabic, occurring in both syllable onsets and rimes; however, the English phoneme /h/ can only occur before vowels and in syllable initial positions. Therefore, /ħ/ and /h/ are contrastive in Arabic where accurate perception of the contrast is crucial for making accurate semantic identification (e.g., *harm* “pyramid” versus *ħarm* “campus”), but such a contrast does not exist in English.

Fourth, the Arabic /ħ/-/h/ contrast is chosen because a number of teachers of Arabic at numerous American universities have encountered more difficulty in identifying productions of /ħ/ by native English speakers than the voiceless glottal phoneme /h/. These speakers are usually unable to perceive the difference between the English speakers’ productions of /h/ and /ħ/, and generally identified them both as /h/ (Shehata, 2007).

Fifth, the Arabic /ħ/-/h/contrast has received no attention in the literature of L2 phonology. There are no previous studies in the L2 literature, to the best of my knowledge, which have investigated the Arabic the /ħ/-/h/ contrast. Taken together, all the aforementioned reasons justify the use of this Arabic contrast in the present research.

2.8 Summary

Several conclusions can be drawn from the findings discussed so far. First, training studies that have examined the effect of talker variability on the learning of novel L2 contrasts have reported conflicting findings. While some studies reported talker variability to facilitate learners' acquisition of novel words for both infants (Spence, Rollins, & Jerger, 2002) and adults (Houston, Jusczyk, & Tager, 1998; Johnson, 1997; Newman & Shannon, 2007; Sumner, 2011; Theodore & Miller, 2010), others found it to be ineffective (Bradlow, Nygaard, & Pisoni, 1999; Goldinger, 1996; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994; Sommers, Kirk, & Pisoni, 1997). A third group of training studies found talker variability to play a minor role in the acquisition of L2 novel contrasts (Hardison, 2003).

Second, talker variability L2 studies mainly used nonlexical tasks in their exploration of the acquisition of L2 novel contrasts, paying little attention to lexical tasks. Third, because controversial results have been reported in the literature regarding the relative difficulty of nonlexical and lexical tasks, the issue needs further research. Fourth, in relation to the acquisition of novel L2 lexical contrasts, most training studies have essentially examined the acquisition of English phonemes by several learners from different L1 backgrounds, such as native Japanese speakers (Logan, Lively, & Pisoni, 1993; Masuda, Norrix, & Green, 2001; Nishi & Kewley-Part, 2007; Pruitt, 1993;

Sheldon & Strange, 1982; Takage & Mann, 1995; Yamada, Stange, Magnusin, Pruit, & Clarke III, 1994), L1 Chinese speakers (Wang, Michelle, Allard, & Joan, 1999), francophone speakers (Jamieson & Morosan, 1986), Spanish and German speakers (Inverson & Evans, 2007), Kikuyu speakers (Streeter & Landauer, 1976), and Korean speakers (Yu & Jamieson, 1993).

It is not clear whether the above-mentioned findings are restricted to the English-speaking data in particular or if they can be extended to include other listener groups from different native language backgrounds. To better understand how training affects L2 phonology, this question needs to be further explored with novel target phoneme contrasts from different native language backgrounds. Obtaining similar findings with other languages would expand the generalizability of this line of investigation and confirm that the findings reported in previous training studies were not confined to particular language populations, but could be extended to include other languages as well. Accordingly, new studies that further investigate this gap in literature are certainly needed.

Given the gap in previous research and based on the robust outcomes of training studies, the present dissertation served as an initial step toward examining the role of talker variability in relation to learners' word recognition and lexical processing of nonnative phoneme contrasts, and posited two main objectives. First, it aimed to contribute to the literature on the acquisition of nonnative contrasting phonemes with the goal of better understanding how talker variability could affect adult L2 learners' identification and lexical processing of unfamiliar phoneme contrasts using a nonlexical task (Experiment 1) and a lexical task (Experiment 2). Second, it investigated the relative effect of task type on the acquisition on nonnative contrasts (by comparing the findings of Experiments 1 and 2).

2.9 Research Questions and Hypotheses

The above-listed primary objectives of the present dissertation address the following research questions and hypotheses:

1. Does perceptual training with multiple talkers, versus a single talker, result in more accurate discrimination of nonnative phoneme contrasts in terms of generalization to novel talkers?

Secondary question, a. Is there more accurate performance with multiple-talker training than with single-talker training on a nonlexical task? (Experiment 1)

Hypothesis 1

Findings from previous training studies with Japanese learners of English, where nonlexical identification tasks were used, demonstrated that multiple-talker training improves listeners' perception of the English /r/ and /l/ (Bradlow et al., 1997; Goldinger et al., 1991; Lively et al., 1993; Logan et al., 1991). Attributes of talkers' voices provide rich input that in turn facilitates learners' identification of that target contrast (Pisoni, 1990). Based on these findings, it is hypothesized that subjects in the multiple-talker training environment in the current study will discriminate the target contrast when it is produced by new talkers, more accurately than subjects in the single-talker training environment.

There are two types of variables in Experiment 1: two independent variables that were item type (target stimuli versus fillers), and training group (single-talker group versus multiple-talker group); and one dependent variable that is the proportion correct at matching given auditory forms and the expected auditory forms. If subjects who listen to the Arabic nonwords (target and filler tokens) introduced by multiple talkers during the word learning task perform more accurately on the nonlexical task than those in the single-talker training group, the benefit of talker variability in the discrimination of the Arabic pharyngeal-glottal contrast by adult native English speakers will be demonstrated.

Secondary question, b. Is there more accurate performance with multiple-talker training than with single-talker training on a lexical task? (Experiment 2)

Hypothesis 2

Findings from previous cross-language studies that used similar lexical tasks (e.g., Cutler et al., 2006) have shown that L2 learners could lexically retain an unfamiliar phoneme contrast after a short exposure. Moreover, the addition of the training session (i.e., the word learning phase) and the practice session (i.e., the criterion test phase) before administering the lexical tasks, which required subjects to access the lexical meaning of test items, might provide subjects with sufficient exposure to the contrast. Consequently, L2 learners' ability to discriminate the contrast and create lexical representations of the newly learned words could be improved. As a result, subjects accurately match the auditory form to the correct picture in the lexical discrimination task. Based on these findings, and taking into consideration the advantage of the multiple-talker training that has been reported in previous L2 training studies (Barcroft & Sommers, 2005; Clopper & Pisoni, 2004; Houston, Jusczyk, & Tager, 1998; Johnson, 1997; Lively, Logan, & Pisoni, 1993; Logan, Lively, & Pisoni, 1991; Rost & McMurray, 2009; Spence, Rollins, & Jerger, 2002), it is expected that subjects in the multiple-talker training group will be able to establish distinctive lexical representations for the newly learned tokens when they are introduced by new talkers, more proficiently than subjects in the single-talker training group.

Experiment 2 includes two main independent variables: item type (target stimuli versus fillers), and training group (single-talker group versus multiple-talker group). The dependent variable, on the other hand, is the proportion correct at matching the auditory form to the correct picture. If subjects in the multiple-talker training group perform more accurately than their counterparts in the single talker-talker training group, it will provide evidence that learners who are trained to distinguish novel tokens spoken by several talkers will have an advantage over the other group. This will be due to the rich input provided by the acoustic characteristics of each talker's voice that are expected to facilitate their accurate discrimination of the newly learned words when spoken by new talkers. Therefore, it will be concluded that training involving variation in talkers' voice can help L2 learners in creating distinct lexical representations of unfamiliar phoneme contrasts.

2. Does task type training (in this case, nonlexical versus lexical) influence learners' ability to discriminate novel L2 phoneme contrasts? (By comparing findings of Experiments 1 & 2)

Hypothesis 3

Following the same methodology as Pater (2003) where the nonlexical and lexical discrimination tasks were made similar by including a word learning phase in the two experiments, it is predicted that subjects will show no difference in performance between the two tasks for the following reasons. First, having received the same amount of the L2 input in terms of quantity and quality (similar auditory and visual stimuli) during both training and criterion test phases, subjects in the two experiments are expected to be able to create similar lexical representations corresponding to the newly learned words and their associated pictures, despite the different types of questions (task demands) they will be asked in the final tests afterwards. Second, having the condition of achieving 90% accuracy in order to be allowed to move to the next task will guarantee subjects' ability to accurately match each of the newly learned words to the correct picture before moving to the final test. Third, the fact that subjects are allowed to repeat training as many times as they like until they get a passing score supports this hypothesis.

There are three independent variables: task type (nonlexical versus lexical), training group (two levels: single talker and multiple talker), and item type (two levels: targets and fillers) as a within-subjects variable. There is one dependent variable, which is subjects' proportion correct (proportion of responses correctly identifying the intended production of the talker in both Experiments 1 and 2). If subjects in the two experiments perform similarly, it will provide evidence that task type has no significant effect when the two tasks are made to be similar.

CHAPTER 3

EXPERIMENT 1: NONLEXICAL DISCRIMINATION, SINGLE-TALKER VS. MULTIPLE-TALKER TRAINING

3.1 Introduction

The primary objective of this experiment was to investigate the role of talker variability in the acquisition of novel L2 phonemes on a nonlexical discrimination task. Therefore, it investigated whether subjects in the multiple-talker training environment could differentiate unfamiliar phoneme contrasts from new talkers more accurately than subjects in the single-talker training environment, in terms of generalization of training to stimuli produced by unfamiliar talkers for subjects in the two training groups.

This experiment was conducted with two groups of native English speakers, in which each group learned a set of Arabic nonword minimal pairs exemplifying the Arabic glottal-pharyngeal fricative contrast in addition to filler items exemplifying the familiar alveolar-postalveolar fricative contrast that is found in both English and Arabic. Subjects were later tested on their ability to differentiate phoneme contrasts that do not exist in their native language, on a nonlexical discrimination task (that did not require lexical access). The two groups differed only in whether subjects listened to one talker or three different talkers producing the stimuli during the word-learning phase.

3.2 Subjects

Participants included 30 native speakers of English (11 males and 19 females) who had no previous significant exposure to the Arabic language. All participants were ages 18 years and higher, living in Utah. They were recruited from undergraduate courses at the University of Utah. Seven subjects received course credit for their voluntary participation; the other 23 subjects received payment ($10 each) for their participation. All subjects reported, via questionnaire, having no speech or hearing problems and no neurological disorders. Participants also reported not being under the influence of any medication that might impact their motor skills. To avoid talker's idiosyncrasies, subjects in this group were randomly assigned to one of the three subgroups: Group1 where subjects listened to stimuli produced only by Talker 1, Group 2 where subjects listened to stimuli produced only by Talker 2, and Group 3 in which subjects listened to stimuli spoken only by Talker 3. See Table 3.1.

Table 3.1 Summary of Training Environments in Experiment 1

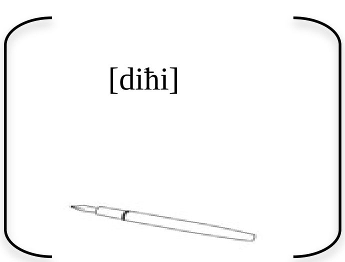
	Word Learning Phase	XAB Nonlexical Task (New Talkers)
Single-Talker Environment	Group1: listened only to Talker 1, Group 2: listened only to Talker 2 Group 3: listened only to Talker 3	The two groups listened to Talker 4, Talker 5 and Talker 6
Multiple-Talker Environment	Subjects listened to Talker 1, Talker 2 and Talker 3	

As shown in Table 3.1, each native English speaker was assigned randomly to one of the two word learning environments: single-talker (7 males and 8 females) and multiple-talker (4 males and 11 females). The duration of learners' participation was approximately 45 minutes.

3.3 Materials

3.3.1 Stimuli

There were two sets of stimuli used in this experiment. The first set included 18 disyllabic Arabic nonwords. These tokens consisted of six minimal pairs contrasting the target Arabic phonemes (i.e., /ħ/ and /h/) in three different positions: initial position (e.g., ħ*aθa-haθa),* intervocalic position (e.g., *di*ħ*i*-dihi), and word-final position (e.g., *ana*ħ *-anah).* The second set included six filler tokens that were three minimal pairs contrasting familiar phonemes found in both English and Arabic as controls in the same vowel environments as the target stimuli: word initial (e.g., *sata*-ʃata), intervocalic position (e.g., *fisi*-*fiʃi*), and in word-final position (e.g., *anas*-*anaʃ*). Figure 3.1 shows an example of auditory and visual forms. Lists of the two sets of target and filler stimuli can be found in Appendix A and Appendix B.



Example Auditory Stimulus

Example Visual Stimulus

Figure 3.1 Example auditory and visual stimuli used in Experiment 1

Each stimulus was randomly assigned to a picture that indicated its false meaning. Using Arabic nonwords and having subjects with no prior exposure to Arabic made it easy to assign any picture to any auditory stimulus.

3.3.2 Talkers

The spoken materials were produced by 6 native male speakers of Arabic (the ‘talkers’ numbered 1-6). They were Egyptians (Egyptian dialect) who were recruited from the University of Utah community. Stimuli were recorded in a sound-attenuated booth using a Marantz PMD 660 recorder and Samson QV microphone in the Speech Acquisition Lab at the University of Utah. Talkers were recorded producing each stimulus in a carrier sentence, “*uridu ?an ?aktubu kalemeta* ________” (I want to write the word **_________**). Table 3.2 provides information about the 6 native Arabic talkers.

Table 3.2 Talker Group: Six Native Arabic (Egyptian) Speakers

	Age	How long learning English	Length of residence in an English speaking country (USA)	Major
Talker 1	24	12	4	Economics
Talker 2	30	16	5	Engineering
Talker 3	26	14	2	Political Science
Talker 4	27	11	9	Physics
Talker 5	23	12	8	Engineering
Talker 6	32	18	10	Education

Talkers were instructed to read the list of 18 Arabic nonwords, which were written in Arabic script at their normal speaking rate. They read the list three times, each time reading the nonwords in a different random order. From these productions, only the second token of each stimulus was extracted from the productions of each talker. The recording process took place during one session.

3.3.3 Procedures

This experiment was designed closely following the methods used by Pater (2003). It was administered in a single session that took place in a sound-attenuated booth. Audio and visual stimuli were introduced using a computer and Sony MDR-7506 headphones that participants used to listen at a comfortable level. Experiment 1 involved three phases: word-learning, criterion test, and nonlexical discrimination test. All phases were shown through the DMDX software that was developed by Forster and Forster (2003).

3.3.3.1 Word-Learning Phase

In the word-learning phase, participants listened to each nonword and saw a picture indicating its meaning. In the single-talker training environment, subjects listened to stimuli produced by a single talker (i.e., either Talker 1 or Talker 2 or Talker 3); however, subjects in the multiple-talker training environment listened to stimuli produced by multiple talkers (i.e. Talker 1, Talker 2, and Talker 3). An example is provided in Figure 3.2.

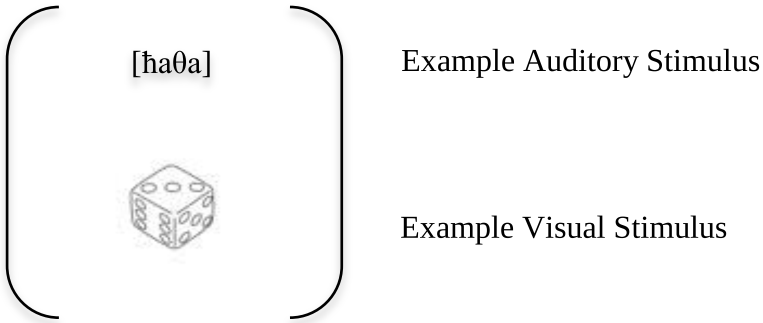


Figure 3.2 Example presentation in the word-learning phase used in Experiment 1

Participants in the two training environments listened to the 18 Arabic nonwords two times per block and each block was presented three times. This resulted in 108 presentations (12 target words + 6 filler words * 2 times per block * 3 blocks) that were presented in random order for each participant in each training environment. There was a 1-second pause between training items that resulted in a 10-minute training session in which subjects listened to the target and filler tokens without any responses and were instructed to learn these tokens and their meanings as well as possible.

3.3.3.2 Criterion Test Phase (Sound-Picture-Picture)

Following the word-learning phase, subjects started the criterion test phase, which was identical for participants in the two training groups; participants were tested on their knowledge of the training stimuli. In the criterion phase, participants were required to associate each word with its correct picture with a 90% or better accuracy before they could move to the XAB nonlexical phase. During this test, they heard a word (X), saw a picture (A), and then saw another picture (B). Subjects' task was to indicate whether the word (X) matched picture (A) or picture (B) by pressing either the right or left shift keys (labeled FIRST and SECOND) on the keyboard. Each word appeared as (X) twice:

one-half was matched with (A) and one-half was matched with (B). Thus, the criterion test included 36 test items (12 target words + 6 filler words * 2 presentations) that were introduced in a different random order for each participant and for each word-learning phase. Participants were given 3 seconds to respond before the program considered their responses incorrect and advanced to the next item.

Participants in the two groups listened to stimuli produced by the same talker(s) they heard in the word-learning phase. This task did not require discrimination of the target contrasts. Participants' responses were considered correct if they matched the given audio word. For example, if the subject heard /*itiħ*/, saw two pictures (of /*sata*/ and /*itiħ*/), and pressed the SECOND key, the response was scored as correct. In order to proceed to the following phase, participants had to score 90% or better on the criterion test phase. Participants, who scored below 90%, began the training phase again. Figure 3.3 displays an example of a criterion test item.

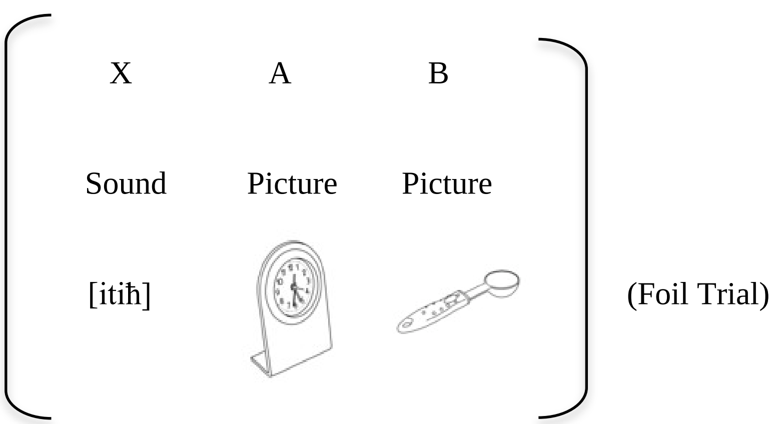


Figure 3.3 Example presentations in the criterion test (Sound-Picture-Picture) used in Experiment 1

There was no specific number of times that participants were allowed to repeat the training phase; however, they could repeat training as many times as they liked until they achieved a passing score. Participants were also provided with feedback on the screen telling them whether or not their performance was accurate enough to pass to the following discrimination test.

3.3.3.3 Nonlexical Discrimination Test Phase (Sound-Sound-Sound)

The third phase was the XAB nonlexical discrimination test phase (sound-sound-sound) where participants were tested on their ability to distinguish the Arabic pharyngeal-glottal minimal pairs. In the XAB Nonlexical Discrimination Test, participants in the two training groups listened to auditory stimuli produced by three unfamiliar talkers: Talker 4, Talker 5, and Talker 6, who had not participated in the word-learning phase for either group. Here, participants heard three nonwords (X, A, and B) and decided whether X was more similar to A or B (e.g., /diħi/-/diħi/-/anah/) by pressing either the right or left shift keys (labeled FIRST and SECOND) on the computer keyboard.

There were 24 trials in the XAB nonlexical discrimination task where A and B were in a minimal pair relationship in addition to 12 foils where A and B were not members of a minimal pair, totaling 36 trials (contrast trials) that were shown in random order. As in the criterion test, there was a 3-second pause after each group of three words for subjects to respond. Figure 3.4 shows an example of the nonlexical test stimuli as presented to subjects in the two training groups.

X	A	B	
Sound	Sound	Sound	
[gaħa]	[gaħa]	[gaha]	(Contrast Trial)

Figure 3.4 Example presentations in the XAB nonlexical discrimination task (Sound-Sound-Sound) used in Experiment 1

The entire experiment lasted for approximately 40 minutes. Subjects in the two groups were instructed to take as many breaks as they liked during both the training and testing phases.

3.4 Coding of Nonlexical Discrimination Test Task

Participants' responses on the nonlexical discrimination task were considered correct only when they matched words as intended by the talkers. For example, if subjects listened to '*haθa-haθa-ħaθa'* in order and then pressed the key labeled FIRST on the keyboard, their answer was counted as correct. Conversely, if they pressed the key labeled SECOND, their answer was considered wrong since it did not match the talker's intended utterance.

3.5 Questionnaire Administration

Participants were instructed to fill out a written questionnaire after they completed the testing phase of this experiment. This questionnaire included questions about each participant's age, gender, and language experience. In addition, there were open-ended questions about participants' impressions of the study and any conscious strategies they might have used in completing the task. A copy of this questionnaire can be found in Appendix C.

CHAPTER 4

EXPERIMENT 2: LEXICAL DISCRIMINATION, SINGLE-TALKER VS. MULTIPLE-TALKER TRAINING

4.1 Introduction

Experiment 2 addressed the second part of the first research question in the current study that examined the influence of training with single and multiple talkers in improving English speakers' performance of the Arabic /ħ/-/h/ contrasts. Here subjects were tested in terms of generalization to new unfamiliar talkers in a lexical discrimination task that required subjects to recall the stored representations of the target contrasts in memory by matching pictures and auditory forms. Therefore, Experiment 2 predominantly tested whether or not subjects were able to store the contrasting phonemes.

4.2 Subjects

The tested population in this experiment included 30 native English speakers with no prior knowledge of Arabic and, who did not participate in Experiment 1. Subjects were recruited from the University of Utah campus and either received course credit or $10 payment for their voluntary participation in Experiment 2. See Table 4.1.

Table 4.1 Summary of Participants in Experiment 2

	XAB Lexical Discrimination Task (New Talkers)
Single-Talker group	15 listeners who were native speakers of English
Multiple-Talker group	15 listeners who were native speakers of English

As Table 4.1 shows, a total of 30 subjects were randomly assigned to 1 of the 2 word learning environments: the single-talker training (4 males and 11 females) and the multiple-talker training (6 males and 9 females). Subjects reported via the given questionnaire, having no speech or hearing problems and no neurological disorders. Subjects' ages were 18 and older and questionnaire data also verified that none of them were under the influence of any medication that might affect their motor skills.

4.3 Materials

4.3.1 Stimuli

The stimuli for the three phases of Experiment 2 were the same as described in Experiment 1, i.e., two sets of stimuli: six target minimal pairs contrasting the Arabic pharyngeal-glottal phonemes in three different phonetic positions, and six fillers contrasting the familiar alveolar-postalveolar fricative contrast (i.e., /s/ and /ʃ/) in the same phonetic environments. Experiment 2 also used the same pictures of Experiment 1 that were assigned to the same nonwords. See the example in Figure 4.1.

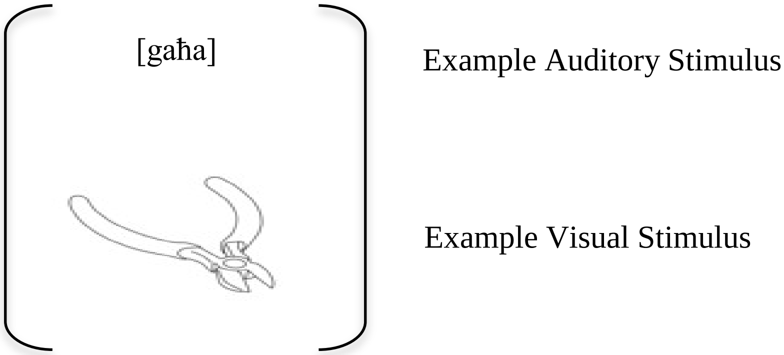


Figure 4.1 Example presentation in the word-learning phase used in Experiment 2

4.3.2 Talkers

Stimuli were produced by the same number of talkers who participated in Experiment 1. For more specific information, please see the detailed description in section 3.3.2.

4.3.3 Procedures

Following the same design in Experiment 1, there were three main phases in this experiment: word learning, criterion test, and final test. Both the word learning phase and criterion tests were similar to those in Experiment 1 where the same auditory and visual representations were used. Again, the word-learning training included 108 tokens (12 target words + 6 filler words * 2 presentations * 3 blocks) and the criterion test included 36 test items (12 target words + 6 filler words * 2 presentations) that were displayed in a different random order for each subject. Participants had to score 90% or better on the criterion test in order to pass to the following test. Those who scored below 90% began the training phase again. Figure 4.2 displays an example of a criterion test item.

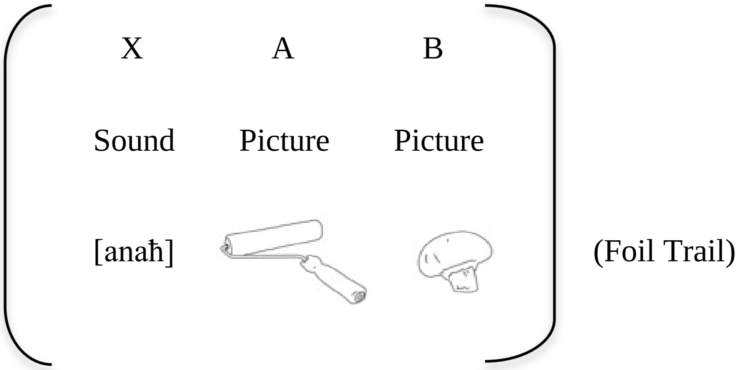


Figure 4.2 Example presentations in the criterion test in Experiment 2

However, the third phase (i.e., final test phase) was different from Experiment 1. Unlike Experiment 1, there was an XAB lexical discrimination test (Sound-Picture-Picture) that was identical to the criterion test where subjects in the two training groups heard a word (X), saw a picture (A), and then saw another picture (B). However, unlike the criterion test, subjects heard tokens from unfamiliar talkers. The subjects' task was to match the word to the correct picture (either A or B) by pressing either the right or left shift keys (labeled FIRST and SECOND) on the keyboard. This task required a discrimination of the target pharyngeal-glottal contrasts where A and B included members of the target minimal pairs (Contrast trial, e.g., /hibi/-/ħibi/).

Similar to the criterion test, there was a 3-second pause after each group of three words for subjects to respond. Subjects' late responses were considered as incorrect by the program that automatically advanced to the next items. In general, the training session continued for approximately 40 minutes. Subjects in the two groups were instructed to take as many breaks as they liked during both training and testing. Figure 4.3 shows an example of stimuli presented in the lexical discrimination test in Experiment 2.

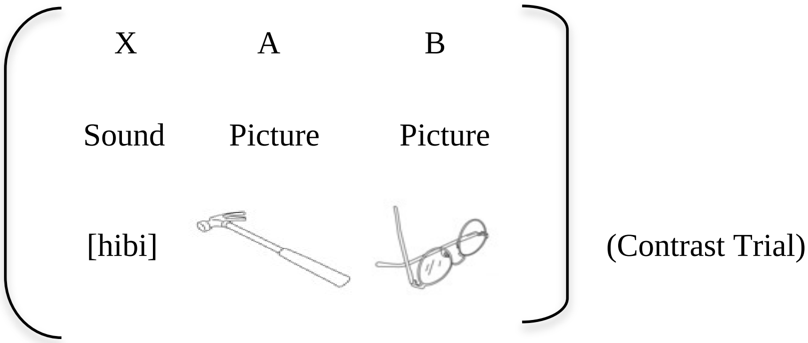


Figure 4.3 Example presentations in the XAB lexical discrimination task (Sound-Picture-Picture) used in Experiment 2

4.4 Coding of Lexical Discrimination Test Task

Subjects’ responses on the lexical discrimination task were coded following the same coding procedures described in Experiment 1. More details about the coding process can be found in section 3.4.

4.5 Questionnaire Administration

Experiment 1 and Experiment 2 used the same questionnaire. Detailed information about this questionnaire is described in more detail in section 3.4, and a copy of this questionnaire can be found in Appendix C.

CHAPTER 5

RESULTS

5.1 Introduction

Recall that the first research question addressed with the present study was: Does training with multiple-talkers, versus a single-talker, result in more accurate discrimination of nonnativeL2 phoneme contrasts in terms of generalization to novel talkers? This question is addressed by considering the results of Experiments 1 and 2 individually, focusing on the difference in performance between subjects in the multiple-talker and the single-talker conditions (sections 5.1 and 5.2). The second research question was: Does task type (in this case, nonlexical versus lexical) influence learners' ability to discriminate novel L2 phoneme contrasts? This question is addressed by comparing subjects' performance on the nonlexical task in Experiment 1 to the lexical task in Experiment 2, in section 5.3 below.

5.2 Experiment 1 (Nonlexical Task) Results

Experiment 1 was designed to examine the impact of single-talker versus multiple-talker training on the recognition of the Arabic pharyngeal-glottal contrasts by learners with no prior experience with Arabic in a nonlexical task. Recall that Hypothesis 1 was that native English speakers in the multiple-talker training group were expected to discriminate the target contrasts on the nonlexical discrimination task more accurately than their counterparts in the single-talker training group.

Proportion correct (proportion of responses correctly identifying the intended production of the talker) was calculated for each participant. The data were submitted to Analysis of Variance, with item type (two levels: targets and fillers) as a within-subjects variable and training group (two levels: single and multiple talker training groups) as a between-subjects variable.

The main effect of training group was also significant (F (1,28)=88.866, p<. 001, partial eta squared = .760), with performance by participants in the multiple-talker training group (.915) more accurate than that of those in the single-talker training group (.671). The effect of item type was significant (F (1,28)=79.646, p<. 001, partial eta squared = .740) with performance on filler items (.911) higher than that on target items (.675). The interaction of item type and training group was also significant (F (1,28)=39.685, p<. 001, partial eta squared =. 586).

Following up on the significant interaction of item type and training group, we will now focus on the results for each item type separately. There was a significant effect of training group on performance on target items (F (1,28)=161.398, p <. 001), with more accurate performance by subjects in the multiple-talker training group (.881) than the single-talker training group (.469). However, the effect of training group on performance on filler items was not significant (F (1,28)=3.564, p =. 069; single talker group: .872, multiple talker group: .950). See Figure 5.1 for visual presentation of these results.

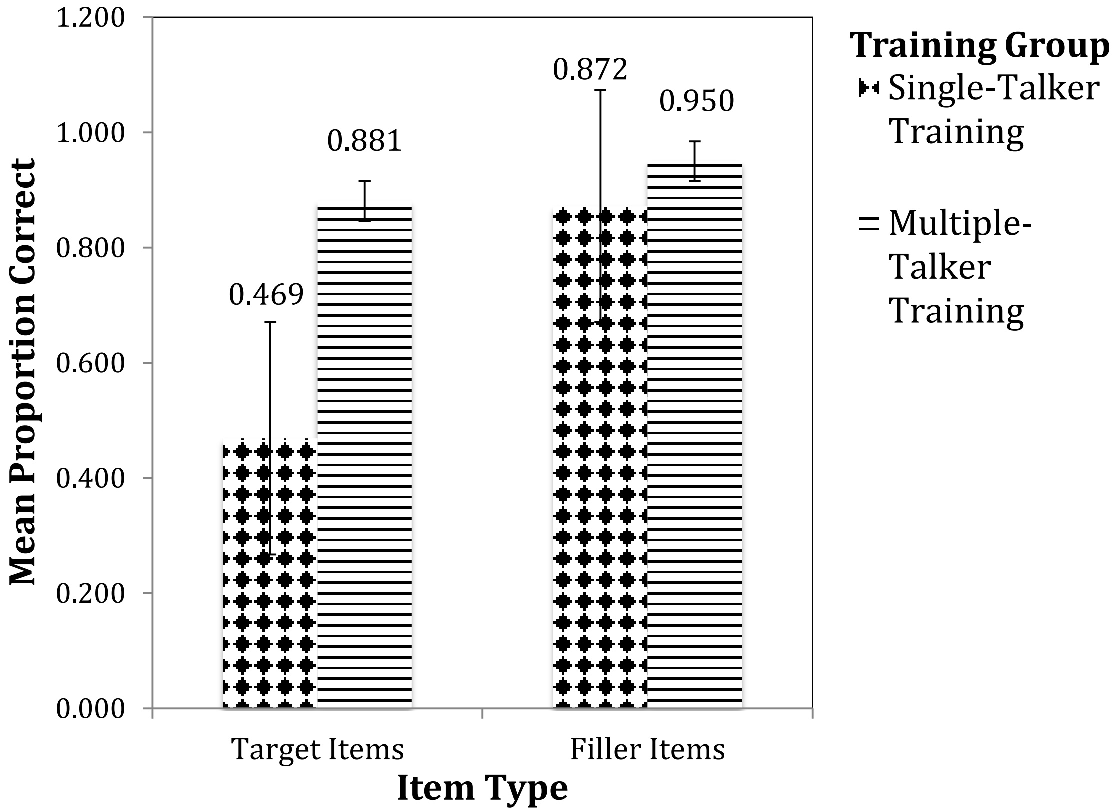


Figure 5.1 Proportion correct for subjects in the two training groups on the nonlexical task; bars represent +/-1 standard error

As Figure 5.1 shows, subjects' performance on filler items, on which it was expected that all subjects would perform well, did not differ significantly between the two groups. Conversely, performance on target items did differ significantly between the groups, and in the expected direction, with subjects in the multiple-talker training group outperforming those in the single-talker training group.

5.3 Experiment 2 (Lexical Task) Results

Experiment 2 was performed to look into the impact of single-talker versus multiple-talker training on the discrimination of the Arabic /ħ/-/h/ phoneme contrasts in a lexical discrimination task. Remember that the hypothesis was, that despite the increasing demands due to the required lexical access in the lexical discrimination task, subjects in the multiple-talker training group would perform more accurately on the lexical discrimination task than subjects in the single-talker training group.

As in Experiment 1, proportion correct (proportion of responses correctly identifying the intended production of the talker) was calculated for each participant. The data were submitted to Analysis of Variance, with item type (two levels: targets and fillers) as a within-subjects variable and training group (two levels: single and multiple talker training groups) as a between-subjects variable. This analysis revealed a significant main effect of training group (F (1,28)= 20.264, p < .001, partial eta squared =. 420), with subjects in the multiple-talker training group (.869) performing more accurately than their counterparts in the single-talker training group (.654).

In addition, the main effect of item type was significant (F (1,28)= 35.598, p < .001, partial eta squared =. 560) with subjects' performance on targets (.676) lower than that on fillers (.847). The interaction of item type and training group was significant as well (F (1,28)= 11.861, p < .001, partial eta squared =. 298).

Following up on the significant interaction of item type and training group, we will now focus on the results for each item type separately. There was a significant difference between the two training groups for target items (F (1,28) = 47.722, p < .001, partial eta squared =. 630), where subjects in the multiple-talker training group performed more accurately (.833) than those in the single-talker training group (.519). On the other hand, the effect of training group on subjects' performance for filler items was not significant (F (1,28) = 3.281, p = .081, partial eta squared = .105). Hence, while performance on filler items, on which it was expected that all subjects would perform well, did not differ significantly, performance on target items did differ significantly between the groups, and in the expected direction, with subjects in the multiple-talker training group outperforming those in the single-talker training group. See Figure 5.2 for visual presentation of these results.

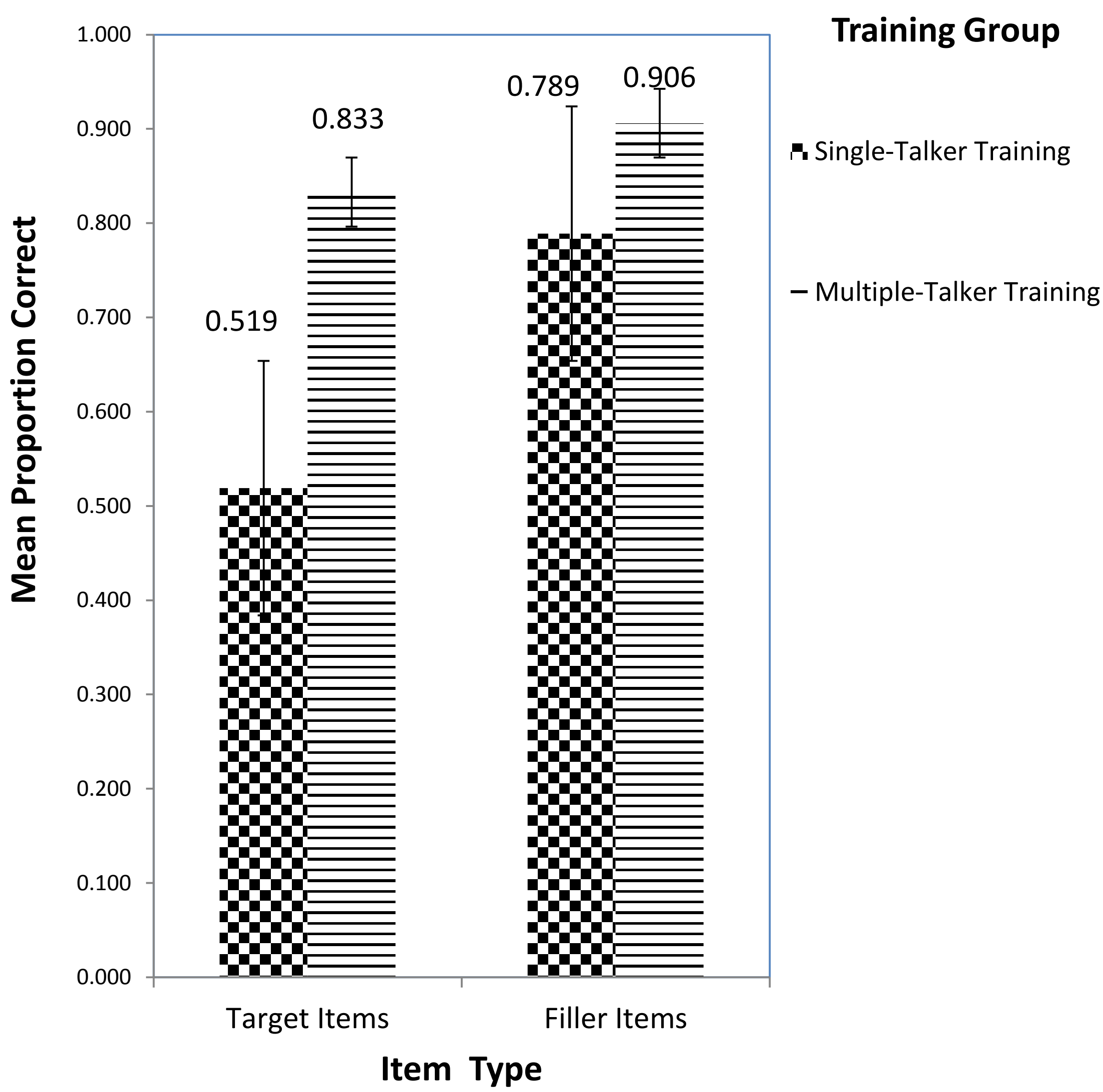


Figure 5.2 Proportion mean correct for subjects in the two training groups on the lexical task; bars represent +/ -1 standard error

5.4 Comparison of Experiment 1 and Experiment 2 Results

As demonstrated above, both Experiments 1 and 2 revealed the expected pattern of results, with subjects in the multiple-talker training conditions outperforming subjects in the single-talker training conditions on target items at test. In order to evaluate the effect of task type—in this case, nonlexical versus lexical—we will now compare the results from Experiments 1 and 2.

An Analysis of Variance with task type (two levels: nonlexical and lexical) and training group (two levels: single talker and multiple talker) as a between-subjects variable and item type (two levels: targets and fillers) as a within-subjects variable was performed. As expected from the results reported above for Experiments 1 and 2 separately, the main effect of item type was significant (F (1,56)=108.965, p <. 001, partial eta squared = .661; target mean: .676; filler mean: .879).

The main effect of training group was also significant (F (1,56)=71.415, p<. 001, partial eta squared = .560; single talker mean: .663, multiple talker mean: .892), as was the interaction of item type and training group (F (1,56)=46.304, p<. 001, partial eta squared = .453).

In contrast, neither the main effect of task type (F (1,56)=1.320, p>. 05, partial eta squared = .023; nonlexical task mean: .793, lexical task mean: .762), nor any of the two-way or three-way interactions involving the task type variable were significant (all p >.05). Together these findings indicate that there was no difference in performance by subjects on the nonlexical versus the lexical tasks. Figure 5.3 visually represents these findings.

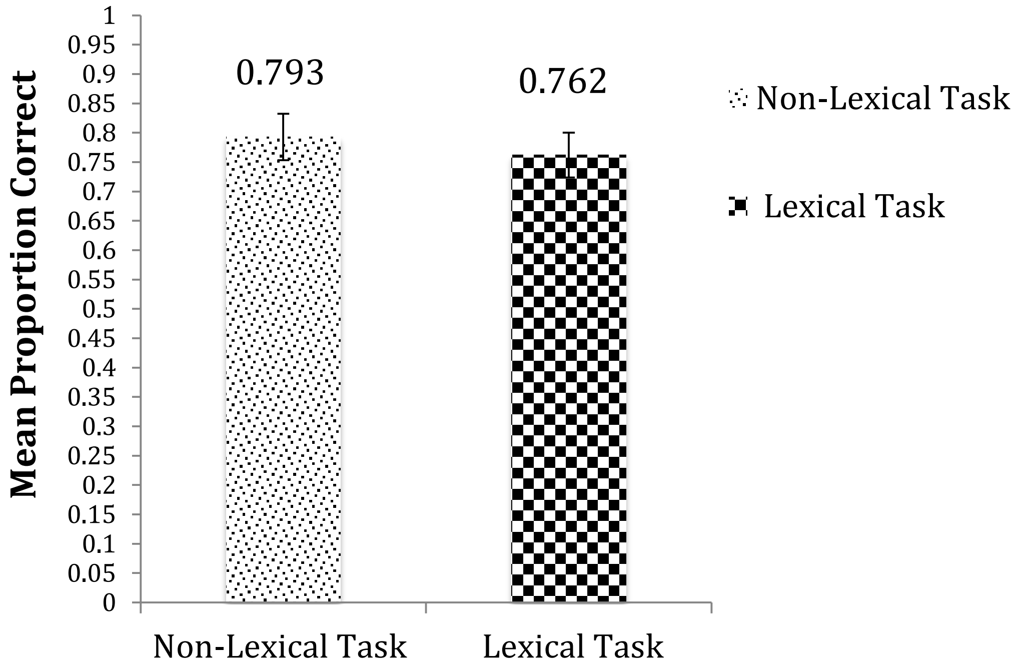


Figure 5.3 Proportion mean correct for subjects in Experiment 1 (nonlexical task) and Experiment 2 (lexical task); bars represent +/ -1 standard error

5.5 Effects of Consonant Position and Vowel Type

It was of interest to investigate whether there was a possible influence of both consonant contrast position and vowel type on subjects' discrimination of target items, due to their significant role displayed in previous research as explained earlier in the literature review in Chapter 2. In order to evaluate that effect, an Analysis of Variance was performed separately on data from the two experiments with task type (two levels: nonlexical and lexical) and training group (two levels: single talker and multiple talker) as a between-subjects variable and consonant contrast position (three levels: initial, intervocalic, and final), and vowel type (two levels: /a/ and /i/) as within-subjects variables. The results of this analysis are summarized in Table 5.1 and visually represented in Figure 5.4.

Table 5.1 Results of Effects of Consonant Contrast Position and Vowel Type for Experiments 1 and 2 Data

	Nonlexical Task (Exp. 1)	Lexical Task (Exp. 2)
Main effect of consonant contrast position	*F* (1,28)= 3.990, *p*< .005, partial eta squared=. 125*	*F* (1,28)= 1.073, *p*> .05, partial eta squared=. 037
Main effect of vowel type	*F* (1,28)= 5.678, *p*< .005, partial eta squared=. 633*	*F* (1,28)= .002, *p*> .05, partial eta squared=. 000
Interaction of consonant contrast position and vowel type	*F* (1,28)=6.026, *p*< .005, partial eta squared=. 866*	*F* (1,28)= 1.615, *p*> .05, partial eta squared=. 327
Interaction of consonant contrast position, vowel type, and the training group	*F* (1,28)=2.109, p>. 05, partial eta squared=. 415	*F* (1,28)=. 642, *p*>. 05, partial eta squared=. 152
Interaction of consonant contrast position and training group	*F* (1,28)=. 203, *p*>. 05, partial eta squared=. 007	*F* (1,28)= .279, *p*>. 05, partial eta squared=. 010
Interaction of vowel type and training group	*F* (1,28)=4.066, *p*>. 05, partial eta squared=. 127	*F* (1,28)= 2.549, *p*>. 05, partial eta squared=. 083

Table 5.1 shows that there was a main effect of vowel type and consonant position in the nonlexical task. There was also a significant interaction of consonant position and vowel type, with the intervocalic (.763) and final positions (.761) being the phonetic environments with the most accurate responses. Conversely, there was not a significant main effect of vowel type and consonant position in the lexical task. In addition, the interaction of vowel type and consonant position was not significant in Experiment 2. Figure 5.4 shows the results of consonant position mean proportion correct by each training group in the two experiments.

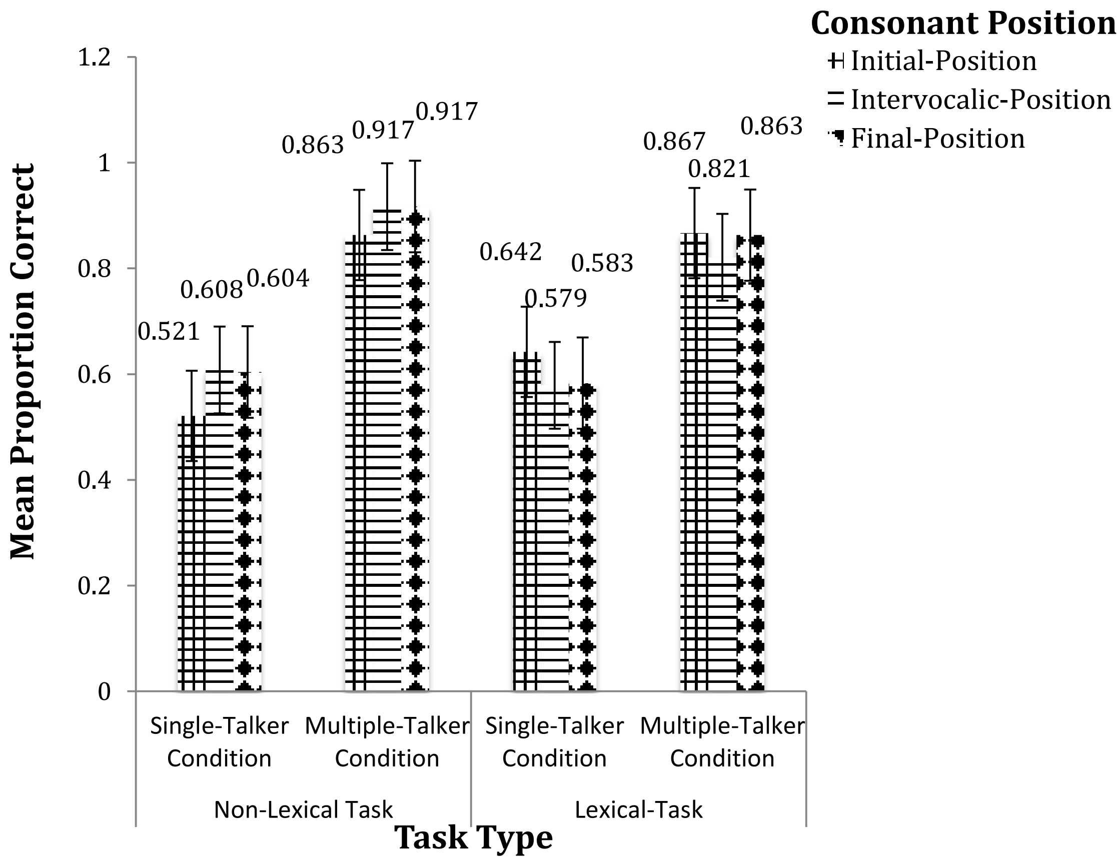


Figure 5.4 Proportion correct on the three consonant positions by groups on both nonlexical (Experiment 1) and lexical tasks (Experiment 2); bars represent +/ -1 standard error

For the nonlexical task, as shown in Figure 5.4, performance of subjects in the multitalker group was high (91% on intervocalic and final positions and 86% on the initial position), whereas performance of subjects in the single-talker group ranged from 52% (on the initial position) to 60% (on intervocalic and final positions) correct. This result indicates that talker variability made little difference in performance. Subjects' performance on the lexical task, on the other hand, was rather interesting. That is, performance of subjects in the two training groups was higher on the word-initial position (86% for the multitalker group and 64% for the single-talker group), than the other two phonetic environments (82% and 86% for the multitalker group and 58% for the single-talker group). Thus, the effect of different consonant position in the nonlexical task was not obtained in the lexical task.

5.6 Summary of Results

Taken together, findings of Experiment 1 and Experiment 2 appear to support the hypotheses of this study, in that performance by subjects in the multiple-talker training groups was significantly more accurately than that of subjects in the single-talker training groups on both lexical (Exp. 2) and nonlexical (Exp. 1) tasks. In other words, subjects who heard the target tokens produced by multiple talkers during training, performed significantly above chance on the test items when they were spoken by new talkers. However, the group of subjects who were exposed to the target items produced by a single native Arabic speaker during training, performed less accurately than the other group when the test items were introduced by unfamiliar native Arabic talkers at test.

The data also indicated that all participants were more accurate at identifying familiar contrastive phonemes from their native language (i.e., /s/ and /ʃ/) than novel ones (i.e., /ħ/ and /h/). Additionally, there was no significant difference in performance by subjects on nonlexical versus lexical tasks. Further discussion of the findings of these two experiments, and their theoretical contribution to our understanding of the discrimination of L2 speech sounds, are addressed in the following chapter in more detail.

CHAPTER 6

DISCUSSION

6.1 Introduction

The present study investigated the effects of talker variability and task type on L2 learners' ability to distinguish and lexically store nonnative contrasting phonemes. This chapter briefly addresses the research questions presented in Chapter 2 with reference to the experiments' findings. Theoretical and pedagogical implications of the study's results are introduced as well.

6.2 Research Question 1

This study involved two main research questions. The first question included two subquestions as follows:

1. Does perceptual training with multiple talkers versus a single talker, result in more accurate discrimination of nonnative phoneme contrasts in terms of generalization to novel talkers?
   a. Is there more accurate performance with multiple-talker training than with single-talker training on a nonlexical task? (Experiment 1)
   b. Is there more accurate performance with multiple-talker training than with single-talker training on a lexical task? (Experiment 2)

2. Does task type training (in this case, nonlexical versus lexical) influence learners' ability to discriminate novel L2 phoneme contrasts? (By comparing findings of Experiments 1 & 2)

6.2.1 Findings of Experiment 1

The purpose of Experiment 1 was to examine whether or not variability in the voice of the talker can positively affect native English speakers' recognition of the Arabic /ħ/-/h/ contrast on a discrimination task that did not require detection of the target contrast. It was hypothesized that participants' accuracy of the Arabic pharyngeal-glottal contrast would increase when the training environment included stimuli produced by multiple talkers.

Findings from Experiment 1 supported this hypothesis. In comparison with the single-talker group, the availability of talker variability in the other training environment did impact subjects' discrimination of the contrasting phonemes. While learners who heard stimuli produced by a single talker were able to discriminate the given contrasts with 67% accuracy (averaged across targets and fillers), learners in the multiple-talker group consistently performed above chance with 92% correct exceeding the accuracy of the other group by 25 percentage points. The interaction between item type and training group was found to be significant as well. The results obtained in Experiment 1 suggest that L2 learners' word recognition can be improved when target items are produced by various voices as compared to repetitions of words in the voice of one speaker. This evidence supports the argument of Logan et al. (1991) and Lively et al. (1993) that variability in talkers plays a positive role in improving learners' discrimination of nonnative contrasting phonemes due to the rich language input that multiple talkers' speech signals provide. Unlike the performance of native Japanese speakers in Logan et al. (1991) and Lively et al. (1993) on the target English /r/-/l/ contrast, whose averaged accuracy scores were 85.9% and 85.6%, respectively, subjects in the multiple-talker training group in this experiment successfully distinguished the Arabic pharyngeal-glottal contrast with 88% accuracy. However, subjects in the single-talker group did not achieve the same success (47%).

One possible explanation for the differences between the two training groups in Experiment 1 may be found in the characteristics of talkers' voices, such as speaking rate, gender, and dialect, which provided extra information to subjects in this training group. That is, the range of variation added by each talker in the multiple-talker training, where distinctive properties of talkers' voices differed from trial to trial, facilitated learners' identification of the contrasting phonemes. In contrast, native English speakers in the single-talker group were not exposed to the same number of speech signals during the training phase. Therefore, their performance was poorer than the multiple-talker group on the nonlexical discrimination task.

In addition, listening to different voices for the first time at the final test added extra difficulty to the task of subjects in the single-talker group because they had to focus on the novel contrast in the auditory forms in order to determine the two similar words in each given auditory input. In addition, they also needed to attend to the new voices whose productions of the newly learned words might sound different from those introduced by the familiar single talker in the previous two phases. As a result, they achieved lower scores than their counterparts in the multiple-talker group.

The present data, on the other hand, argue against the view that considers variation in talkers' voices to play a negative effect in language learning because it demands additional time to store both the linguistic (including the morphological, phonological, syntactic components) and indexical (characteristics of talker's voice) information of speech signals in learners' lexicon (Martin, Mullennix, Pisoni, & Sommers, 1989; Mullennix & Pisoni, 1990; Mullennix, Pisoni, & Martin, 1989; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994; Sommers, Kirk, & Pisoni, 1997). Instead, talker variability was found here to enhance listeners' recognition of the unfamiliar contrast.

6.2.2 Findings of Experiment 2

Experiment 2 investigated the possible contribution of talker variability to the training of adult native English speakers to create lexical representations for the novel Arabic phoneme contrast on a lexical task that required discrimination of the contrasting phoneme. Recall that the hypothesis of this experiment was that subjects' accuracy of the target phonemes would increase under the multiple-talker group compared to the single-talker group.

Findings of Experiment 2 supported the second hypothesis of the present study. Participants in the multiple-talker training group were significantly more accurate at discriminating the contrasting phonemes than participants in the single-talker training group. Not only were subjects able to detect the difference between the target tokens based on the speech sound contrast, but they also successfully exploited this phoneme contrast to discriminate the meaning of words in the lexical identification task. For example, realizing that the two tokens *diħi* and *dihi* refer to two different lexical items (i.e., pen and a paper clip, respectively) provided subjects in the multiple-talker group the adequate information to detect the difference between their middle consonant phonemes, /ħ/ and /h/, and that knowledge accordingly helped them establish phonetic categories of the target contrasts (with 83% accuracy). However, subjects in the single-talker training environment showed a relatively lower discrimination performance (with 52% accuracy). Despite the difficulty of the XAB lexical task that was reported by Pater (2003), subjects in the multiple-talker group were able to create lexical representations for the Arabic minimal pairs in their mental lexicon. This result provides additional evidence supporting the positive role of talker variability in the acquisition of one of the difficult Arabic consonant contrasts (i.e., /ħ/-/h/) that learners of Arabic often find challenging to acquire. This can provide more robust results regarding the beneficial role of talker variability in L2 acquisition.

In line with findings from Experiment 1, moreover, subjects in the two training groups in Experiment 2 were more accurate at detecting consonant contrasts from their native language (i.e., /s/-/ʃ/ in fillers (with 85% accuracy) than novel contrasting phoneme (i.e., /ħ/-/h/) in targets (with 68% accuracy). While subjects in the single-talker group identified the familiar phoneme contrast with 79% accuracy, learners in the multiple-talker groups were able to detect these two familiar phonemes with 91% accuracy.

Considered together, findings from the two experiments provide further evidence for the benefit of the high-variability training paradigm (Logan, Lively, & Pisoni, 1991). According to this type of training study, the availability of various sources of variability (e.g., stimuli, talkers, and phoneme environments) while learning unfamiliar contrasting phonemes facilitated the learning process of the Arabic pharyngeal-glottal contrast, resulting in an increase of learners' response accuracy on the given tests. This is consistent with those previous studies that have reported the ability of learners to establish lexical representations for nonnative phoneme contrasts (Cutler, Weber, & Otake, 2006; Hayes-Harb & Masuda, 2008; Weber & Cutler, 2004). Although participants in the present study knew no Arabic, they were able to accurately distinguish the target phonemes when they were produced by several talkers. Overall, performance of native English participants in the multiple-talker training groups on the target tokens was above 85% in the two experiments.

Contrasting the proposal of Nygaard and Pisoni (1998) and Nygaard, Sommers, and Pisoni (1994) that found familiar talkers' voice information to be used as an aid to word recognition, results of the present study exhibited variability in talker's voice as important information that is used in mapping the phonetic categories, and consequently cannot be ignored. The difference between the results of the three studies can probably be attributed to the differences in the design of each. Unlike the present study, Nygaard and colleagues' studies presented the target stimuli, i.e., isolated word stimuli in Nygaard et al. (1994), and isolated word and utterance stimuli in Nygaard and Pisoni (1998) at different signal noise ratios to subjects who were instructed to write down the stimuli they heard. One more reason for this difference may be related to the nature of stimuli in each study (i.e., real English words and sentences in Nygaard and colleagues studies versus Arabic /ħ/-/h/ minimal pairs in the present study).

6.3 Research Question 2

The second research question was as follows: Does task type (in this case, nonlexical versus lexical) influence learners' ability to discriminate novel L2 phoneme contrasts?

To contribute to previous research that looked at whether learners' discrimination of novel contrasts differs depending on task type, one of the main objectives of the present study was to examine the possible task type effect on L2 learners' acquisition of novel contrasts. Two different XAB discrimination tasks were implemented at test: nonlexical and lexical tasks. Due to the different demands of each task (e.g., a lexical task typically requires lexical access to the memory that is not required by the nonlexical task), subjects' performance on the two tasks might be expected to differ. Recall that mixed results were reported in the literature regarding this issue. While some studies have provided evidence that subjects' performance differs on the two types (Curtin et al., 1998; Hayes-Harb, 2007; Hayes-Harb & Masuda, 2008), others found no difference in performance, specifically, when the two tasks were made to be similar (Pater, 2003).

6.4 Comparison of Findings of Experiments 1 and 2

Extending the research done by Pater (2003), the second research question asked whether or not task type affects learners' acquisition of novel contrasts. Recall that Hypothesis 3 expected no difference in performance by participants on nonlexical and lexical tasks. Findings confirmed this hypothesis. Figure 6.1 below visualizes this finding.

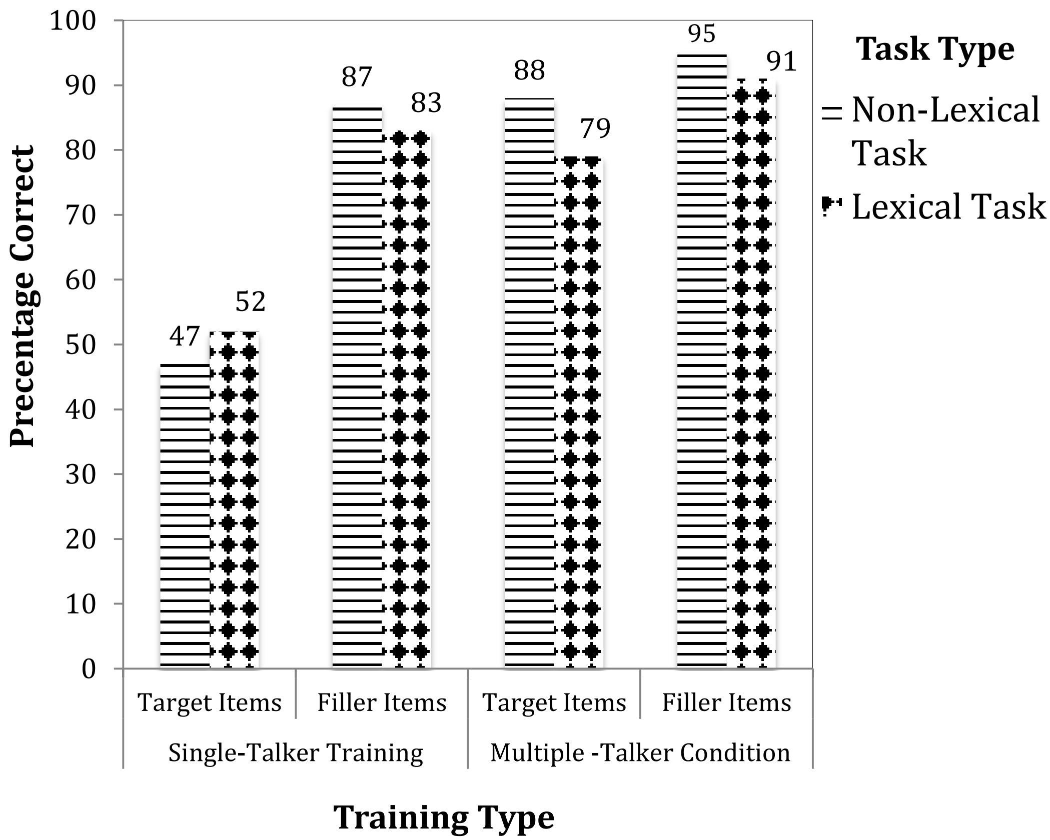


Figure 6.1 Differences in performance by task (Experiment 1 vs. Experiment 2)

Comparing subjects' performance in the two experiments demonstrated no significant difference between subjects' performance on the nonlexical task (92% correct) versus the lexical task (87% correct), despite the different demands of each task that might result in different results. This finding, shown in Figure 6.1, contradicts findings from previous research, such as Curtin, Goad, and Pater (1998), which found native English speakers with no Thai background to be more accurate at distinguishing Thai aspiration contrasts in a nonlexical task than a lexical task. The divergent results of the two studies may be attributed to the difference in the nature of the lexical task used in each case. While Curtain and colleagues (1998) used a picture selection task in which subjects were asked to decide which of the three pictures (two of them were pictures of words in a minimal pair relationship) they saw on the screen matched the word they heard, by pressing the number of the picture on a computer keyboard, the present study used the XAB lexical sound -picture-picture (SPP) task where subjects decided which of the two given pictures matched the word they heard. Although the two lexical tasks were preceded by training phases, they had different requirements that can explain the different findings. It can be argued that the naming picture task is more demanding for listeners since it included three pictures that can present a higher level of complexity than the SPP task. Therefore, subjects might find this task to be more challenging than the nonlexical task given in Curtain and colleagues (1998) where they were able to accurately discriminate the target Thai contrasts.

Conversely, this finding is consistent with Pater (2003) whose findings showed no difference between subjects' performance on the two XAB tasks (78% correct on both tasks). One possible reason for this finding maybe due to the similar L2 input that subjects in the two experiments received. Both experiments, as pointed out earlier, included the same word-learning phase in which subjects learned the auditory and visual stimuli and the same criterion test that checked subjects' learning of the novel words. Having the same information, whether introduced by one or multiple talkers, provided subjects with the same input that presumably resulted in mapping stimuli consistently during the experiments' two different stages (Shiffrin & Schneider, 1997) regardless of the different demands of each of the tasks that subjects performed afterwards. In other words, it can be claimed that subjects in the two experiments started the final discrimination task, nonlexical or lexical, with the same mental representations of the newly learned words. Therefore, the different demands of the tasks did not influence their performance.

Even though the two studies have displayed similar findings, they did report different results concerning the performance of subjects on the SPP task. Unlike subjects in Pater's study, whose performance was at chance on the lexical SPP task, performance of subjects in the present study was relatively higher with 67.6% accuracy. Because the two SPP tasks were similar in the two studies, i.e., they included the same number of target tokens (i.e., 18 tokens) spoken by multiple talkers and simultaneously presented with pictures, this difference is perplexing.

Despite the fact that the two contrasting features in the two studies, i.e., the Arabic pharyngeal and the Thai aspiration, are not found in English, one might claim that the Arabic pharyngeal-glottal contrast, which was reported by learners of Arabic to be difficult to learn in comparison with other Arabic consonant contrasts (Shehata, 2007), is relatively easier for native English speakers to learn than the Thai aspirated-unaspirated consonant contrast. However, this claim needs to be further explored before any conclusions can be made.

6.5 Talker Variability and Word Recognition

Little has been understood about the role of a talker' voice characteristics in mapping the spoken L2 words to meaningful units. While a number of infant training studies examined the influence of talker variability on learners' word recognition ability that represents the intermediate stage between the lower level of perception and the high cognitive processes of interpretation, adult L2 training studies have not paid due attention to this issue. Instead, previous research has looked at the role of talker variability in perceiving L2 novel phoneme contrasts belonging to different phonetic categories (Logan et al., 1991). Therefore, the present study further extends this investigation to L2 word learning and word recognition.

According to traditional views of spoken word recognition, variation in voice is problematic and, therefore, speech produced by different talkers should be normalized to reduce variability. For example, Halle (1985) claimed, “When we learn a new word we particularly never remember most of the silent acoustic properties that have been present in the signal that struck our ears” (Halle, 1995, p. 101). As a result, earlier research on word recognition tended to control for the different sources of variability, i.e., stimuli, talkers and phonetic environments (Byrd, 1992). However, recent adult L1 studies, as explained in section 2.4.2.1, displayed variability to be a useful aid for listeners (Goldinger, Pisoni, & Logan, 1991; Mullennix & Pisoni, 1990; Mullennix, Pisoni, & Martin, 1989; Plameri, Goldinger, & Pisoni, 1993; Pisoni, 1990; Pisoni, 1992). Therefore, recent approaches to word recognition paid much attention to variability in general and talker variability in particular. Under the assumptions of recent approaches, both linguistic information and talkers’ voice information are stored in learners’ long-term memory (Goldinger et al. 199; Martin et al., 1989; Palmeri et al., 1993) where they are retained for some time after learning (Goldinger, 1996). This is supported by multiple-talker subjects’ high accuracy rates on the target words in the present study.

The influence of talker variability has been explained using the framework of the new exemplar-based approach (Goldinger, 1998; Pisoni, 1997). A main assumption of this approach is that acoustic characteristics of target tokens produced by different talkers, which include indexical (i.e., information about talker's gender, age, dialect, social class, and speaking rate) and phonetic information, are stored in learners' mental lexicons, resulting in facilitating recognition of novel representations of these target words when they are produced by new talkers. As Pisoni said, "listeners encode speech signals in multiple ways along many perceptual dimensions, and the nervous system apparently preserves these perceptual details much more reliably than researchers have believed in the past" (1997, p. 71). This has been confirmed by the better performance of native Japanese speakers in the multiple-talker group on the English /r/-/l/ contrast compared to the single-talker group in Logan et al. (1991). In the case of native English speakers in the present study, the number of representations of each target token (i.e., three instances for each word) acquired by subjects in the multiple-talker groups during training helped them in identifying the same tokens when produced by novel talkers during the test phase. In contrast, subjects in the single-talk groups stored fewer representations (one instance for each token) and these, therefore, did not help them in distinguishing the novel productions of the new talkers. This can explain the accurate performance of the multiple-talker groups.

It is also noteworthy that the influence of the consonant phonetic environment varied in the two experiments. Namely, while subjects' performance significantly differed according to the phonetic environment where /ħ/ and /h/ occurred on the nonlexical task, that difference was found to be insignificant on the lexical task. Overall, subjects performed more accurately on both intervocalic and final-word environments (.762) than the initial-word environment (.692) in Experiment 1. Such a finding has been reported with other phoneme contrasts (e.g., /r/ and /l/) in the literature (Dissosway-Huff, Port, & Pisoni, 1982; Lively et al., 1992; Logan et al., 1991; Sheldon & Strange, 1982). This further supports the idea that different acoustic characteristics distinguishing contrastive phonemes are enhanced in the intervocalic and final positions, but are reduced in initial position, consequently influencing subjects' performance (Sheldon & Strange, 1982). In contrast, subjects' performance in Experiment 2 showed a different pattern. That is, subjects in the two training groups performed more accurately on the initial-word environment (.755) than the other two phonetic environments (.711). However, the different performance of subjects on the three phonetic environments in Experiment 2 was not significant. This is a puzzling finding that demonstrates the need for further research on this topic.

6.6 Theoretical and Pedagogical Implications

Generally, findings of the present study are beneficial for both researchers and teachers alike. They draw their attention to the effective role that talker variability can play in other real life contexts and raise interesting questions such as, what may be the implications related to adults' social setting and listening to different talkers? Do adult L2 learners who live in a foreign country and are exposed to the target language through participating in various social activities have advantage over others who lack the same amount of exposure to the target language (and allegedly contact fewer talkers of the target language)? How are novel features on L2 phonemes initially stored? And how are they transferred from learners' working memory into their long-term memory? How is indexical and linguistic information stored in learners' lexicons (same or separate units)? These questions need to be considered, as their answers can possibly help us better see the big picture of speech perception development with reference to variability in talkers, and consequently improve our understanding of this issue.

For teachers, findings from the two experiments have significant practical implications in the area of L2 instruction. First, they shed some light on the significance of using rich acoustic presentation formats in the introduced input, specifically, in novel L2 sound learning. This in turn leads to improvement in teaching methods and techniques that can enhance the given auditory input. Second, results draw L2 teachers' attention to consider effects of task type (specifically those that are not similar, on learners' acquisition of the target language) when designing or choosing tasks to use in their classrooms. In this case, it is recommended to use "less controlled tasks," similar to the criterion tasks used in the present studies that can better prepare learners for the demands of the following tasks after checking their mastery of the target sounds.

Finally, teachers and designers of L2 materials are recommended to take into consideration the use of novel phonetic categories in various phonetic environments since these are shown to be an effective variable in the acquisition of nonnative contrasts.

CHAPTER 7

CONCLUSION

The present study examined the relevant contributions of both talker variability and task type to L2 learners' ability to correctly identify and store unfamiliar phoneme contrasts in the lexicon. Talker variability was manipulated in two different training environments: a single-talker training where subjects heard the target tokens produced by a single talker; and a multiple-talker training in which subjects had three different talkers reading the same target tokens. It was hypothesized that if subjects in the multiple-talker training environment could perform more accurately on both nonlexical and lexical tasks than those in the single-talker training environment, then training that considered acoustic variability must be beneficial to L2 learners, facilitating their learning of novel phonemes. Moreover, task types were hypothesized not to influence subjects' performance.

Results from the two experiments supported the study hypotheses. As predicted, the overall accuracy of learners in the multiple-talker training groups in the two experiments was greater than that of the other two single-talker training groups. Even though subjects' improvement in identification performance varied across tasks (e.g., nonlexical versus lexical tasks), this difference was found to be insignificant.

Taken as a whole, results of the present study addressed some significant issues in the field of L2 speech in general, and word recognition in particular. First, it provided evidence that native English speakers were able to discriminate the Arabic target contrast. The performance of native English speakers on target items was above 67% on both lexical and nonlexical tasks.

Second, the findings successfully displayed that variability in a talker's voice could yield a significant improvement in learners' identification and lexical processing of the Arabic pharyngeal-glottal phoneme contrast that is not contrastive in their native language. Third, findings also introduced additional evidence that when training incorporated high-variability paradigms in terms of stimuli, talkers, and phonetic environments, it not only influenced learners' ability for discrimination ability of novel contrasting phonemes, but their higher level of language processing (i.e., lexical access) as well, as opposed to early training studies that lacked this advantage. Recall that subjects in the two experiments self-reported via the given questionnaire having no prior knowledge of Arabic, yet results displayed a significantly better improvement in word recognition ability for the two multiple-talker training groups than the single-talker training groups.

Fourth, while results displayed native English speakers to be more accurate at distinguishing contrasting phonemes from their native language than nonnative phoneme contrasts, storing the novel phonemes to the lexicon was not found to increase difficulty in discriminating the novel contrasting phonemes. Fifth, unlike findings of Experiment 1 that revealed L2 learners' performance to be influenced by the phonetic environment where the target contrasts occurred, results of Experiment 2 indicated no significant influence of the phonetic environment on learners' lexical processing of the target contrasts. This means that the demands of each task did not influence subjects' general performance, but did have an effect on subjects' performance on the different phonetic environments, which can be further explored in future research. Finally, learners' performance on the nonlexical XAB task did not differ from that of their counterparts on the lexical XAB identification task.

Generally, the present study raised various interesting questions about the development of lexical representations in memory. Further research is still needed to examine the robustness of the present findings and to further answer these significant questions.

CHAPTER 8

LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

In light of the abovementioned findings of this dissertation, some directions for possible follow-ups to the present experiments can be recommended. More data are still needed to test the robustness of the current findings, with more learner groups and/or new stimuli. Furthermore, the present experiments may be replicated, but using other salient Arabic contrasts including both consonants and vowels that have not previously been examined, to decide if they follow the same pattern reported in the current data. This exploration can enrich word recognition investigation in particular and L2 speech research in general.

Even though this study used a high-variability training paradigm to examine learners' ability to generalize training to stimuli produced by novel talkers, it did not investigate subjects' long-term retention of the newly learned Arabic nonwords, thus needing to be considered in follow-up studies. It might be the case that the training used in this study has more detectable impact on L2 learners' performance when subjects are examined over a longer span of time. This would provide robust results about the influence of training that, in turn, can help learners develop distinct categories for target sound contrasts. In addition, replicating this study with production tests is another direction for future research where participants are presented with target tokens and instructed to name each picture when it is presented. Also, doing acoustic analyses of the production data in the replicated study could help in identifying ways in which subjects distinguish the Arabic pharyngeal-glottal phoneme contrast.

Future investigations need to particularly address the question of whether findings from the present study, which focused on segmental-word level using minimal pairs of nonwords spoken in isolation, can be obtained in other training settings where word identification tasks include real words produced in connected speech. Future work can also examine whether or not familiarity with the target language enables L2 learners to create target-like lexical representations of novel contrasting phonemes and consequently eliminate the nonnativelike representations in their L2 speech.

It is also recommended that future experimental investigations consider comparing subjects' performance on different lexical tasks (i.e., sound-picture-picture versus picture-sound-sound) that can further enrich our understanding of learners' word recognition ability. Further research might explore how correct features of contrasting phonemes develop over time within individual learners, through longitudinal data using sentence stimuli that were used in a very few number of studies (e.g., Nygaard & Pisoni, 1995), whereas the bulk of previous studies mainly used isolated words (nonsense and real words). All these proposed directions for future research will result in a better understanding of adult L2 learners' word recognition ability and could yield significant suggestions for instructors and curriculum developers regarding the acquisition of nonnative phonetic categories.

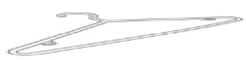
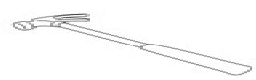
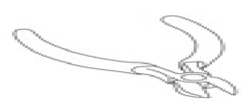
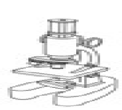
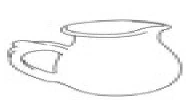
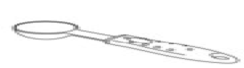
Finally, a multidisciplinary approach, involving several disciplines such as cognitive psychology, phonology, and SLA, which integrates auditory and visual input and explores L2 speech processing beyond the segmental level, is another gap that deserves further consideration in future work. This approach is recommended to include rich input that is characterized with variability in talkers that has proved to be advantageous in the acquisition of novel L2 contrasting phonemes. This type of research is hoped to provide evidence for the most successful L2 pedagogy.

APPENDIX A

A LIST OF TARGET ITEMS USED IN EXPERIMENT 1

AND EXPERIMENT 2 WITH THEIR

ACCOMPANYING PICTURES

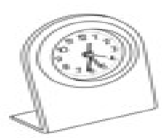
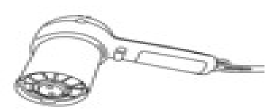
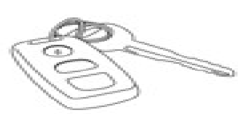
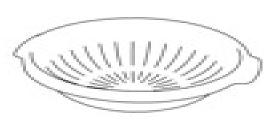
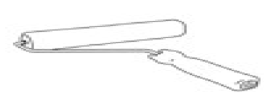
Stimuli / Consonant Position	Auditory Form	Visual Form	Auditory Form	Visual Form
Initial Position	[haθa]		[ħaθa]	
	[hibi]		[ħibi]	
Intervocalic Position	[gaha]		[gaħa]	
	[dihi]		[diħi]	
Final Position	[anah]		[anaħ]	
	[itih]		[itiħ]	

APPENDIX B

A LIST OF FILLER ITEMS USED IN EXPERIMENT 1

AND EXPERIMENT 2 WITH THEIR

ACCOMPANYING PICTURES

Stimuli / Consonant Position	Auditory Form	Visual Form	Auditory Form	Visual Form
Initial Position	[sata]		[ʃata]	
Intervocalic Position	[fisi]		[fiʃi]	
Final Position	[anas]		[anaʃ]	

APPENDIX C

LANGUAGE BACKGROUND QUESTIONNAIRE[2]

ID ______________________________
Please do not write in this box

Date: ______________________________

1. Gender: (circle one) Female Male

2. Age: ___________

3. Do you have any speech, language, hearing, or neurological disorders? Yes No
   If yes, please describe on the back of this page.

4. Are you taking any medications that may affect your motor skills? Yes No
   If yes, please describe on the back of this page.

5. What language do you speak with your parents? ______________________________

6. What language do you consider to be your native language? ____________________

7. What other language/s do you know and how well do you know it/them?
   Language 1: _________________ (circle one) Basic Conversational Fluent
   Language 2: _________________ (circle one) Basic Conversational Fluent
   Language 3: _________________ (circle one) Basic Conversational Fluent
   If you speak additional languages, please provide information about them on the back of this page.

8. What comments do you have about this task?

9. Was the task difficult/easy? What made it difficult?

---

[2] Adapted from the materials of the Speech Lab Acquisition, Linguistics Department at the University of Utah

10. Were you sure about most of your judgments, or did you feel like you were guessing?

11. What do you think we are testing with this task?

REFERENCES

Allaga-Garcia, C. (2007). The role of phonetic training in L2 speech learning, *Proceedings of the Phonetics Teaching & Learning Conference.* London: University College London, August 24-26.

Aliaga-García, C., & Mora, J. C. (2009). Assessing the effects of phonetic training on L2 sound perception and production. A M. A. Watkins, A. S. Rauber, & B. O. Baptista (Eds.), *Recent research in second language phonetics/phonology: perception and production* (pp. 2-31). Newcastle upon Tyne: Cambridge Scholars Publishing.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, *32*, 233–250.

Asher, I. I., & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal, 53*, 334-41.

Barker, B. A., & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, *94*, B45-B53.

Baker, W., & Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: the role of individual differences. *IRAL*, *44*, 3, 231-250.

Baker, W., Trofimovich, P., Mack, M., & Flege, J. E. (2002). The effect of perceived phonetic similarity on nonnative sound learning by children and adults. *Proceedings of the Annual Boston University Conference on Language Development*, *26*(1), 36-47.

Barcroft, J., & Sommers, M. S. (2005). Effects of acoustic variability on second language vocabulary learning. *Studies in Second Language Acquisition*, *27*, 387–414.

Bell, T. S., & Wilson, R. H. (2001). Sentence recognition materials based on frequency of word use and lexical confusability. *Journal of the American Academy of Audiology*, *12*, 514-522.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L. J., & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General*, *117*, 21-33.

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), S*peech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Timonium, MD: York Press.

Best, C. T., McRoberts, G.W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345-360.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, *20*(3), 305-330.

Boatman, D. (1990). The perception and production of a second-language contrast by adult learners of French. In New Sounds 90: *Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-Language Speech* (University of Amsterdam, 9-12 April 1990), Ed. J. Leather and A. James. University of Amsterdam.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to nonnative speech. *Cognition*, *106*(2), 707–729.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101*(4), 2299-2310.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words, *Perception & Psychophys*, *61*, 206–219.

Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and nonnative listeners: talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America*, *106*(4), 2074-2085.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, *61*(5), 977-85.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar position. *Journal of the Acoustical Society of America*, *117*, 3890–3901.

Cartell, T . D. (1984). Contributions of fundamental frequency, formant spacing, and glottal waveform to talker identification. *Res. Speech Percept. Tech. Rep*. No. *5* (Indiana Univ., Bloomington, IN).

Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation*.

Cambridge: Cambridge University Press.

Clopper, C., & Pisoni, D. (2004). Effects of talker variability on perceptual learning of dialects. *Language and Speech, 47*(3), 207-239.

Craik, F., & Kirsner, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology, 26*, 274-284.

Creelman, C. D. (1957). Case of the unknown talker. *Journal of the Acoustical Society of America*, *29*, 655.

Curtin, S., Goad, H., & Pater, J. V. (1998). Phonological transfer and levels of representation: The perceptual acquisition of Thai voice and aspiration by English and French speakers. *Second Language Research*, *14*(4), 389–405.

Cutler, A. (1989). Auditory lexical access: where do we start? In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 342-355). Massachusetts Institute of Technology Press: Cambridge MA.

Cutler, A., & Broersma, M. (2005). Phonetic precision in listening. In W. J. Hardcastle & J. M. Beck (Eds.), *A figure of speech: A festschrift for John Laver* (pp. 63-91). Mahwah, NJ: Erlbaum.

Cutler, A., & Otake, T. (2004). Pseudo-homophony in nonnative listening. *Journal of the Acoustical Society of America*, 115, 2392.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second- language listening. *Journal of Phonetics*, *34*, 269–284.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mother’s voices. *Science, 20X*, 1174-l 116.

Dissosway-Huff, P., Port, F. R., & Pisoni, B. D. (1982). Context effects in the perception of English /r/ and /I/ by Japanese. *Research on Speech Perception Progress Report, 8*, 277-288. Indiana University.

Elman, J. L., & McClelland, J. L. (1984). The interactive activation model of speech perception. In N. Lass (Ed.), *Language and speech* (pp. 337-374). New York: Academic.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in second language acquisition*, *26*(4), 551-585.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, *36*, 345-360.

Fant, G. (1973). Speech Sounds, and Features. Cambridge, MA: MIT Press.

Farnsworth, L. M., & Mullenix, J. W. (1995). *The effects of talker variability across CV and VC environments*. Paper presented at the 129th meeting of the Acoustical Societal of America, Washington, DC.

Flege, J. E. (1984). The effect of linguistic experience on Arabs' perception of the English *Isl* vs. *Izl* contrast. *Folia Linguistica, 18(1-2),* 117-138.

Flege, J. E. **(**1987**)**. The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15,* 47–65.

Flege, J. E. (1988). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human communication and its disorders* (pp. 224–401). Norwood, NJ: Ablex.

Flege, J. E. (1989). Chinese subjects' perception of the word-final English /t/–/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America, 86,* 1684–1697.

Flege, J. E. (1995). Two procedures for training a novel second language phonetic contrast. *Applied Psycholinguistics, 16,* 425–442.

Flege, J. E., Bohn, S., & Jang, S. (1997). Effects of experience on nonnative speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437-470.

Flege, J.E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition, 23,* 527-552.

Flege, J. E., & MacKay, I. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26,* 1-34.

Flege, J. E., MacKay, I. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America, 106*(5), 2973-2987.

Flege, J., Munro, M., & MacKay, I. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication, 16,* 1-26.

Flege, J., Takagi, N., & Mann, V. (1995**)**. Japanese adults can learn to produce English /[/ and /l/ accurately. *Lang. Speech, 38,* 25-56.

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *Journal of the Acoustical Society of America*, *99*, 1161—1173.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computer*, *35*, 116-124.

Frauenfelder, U. H., Baayen, R. H., & Hellwig, F. M. (1993). Neighborhood density and frequency across languages and modalities. *Journal of Memory and Language*, *32*, 781- 804.

Gass, S., & Evangeline, M. V. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*(1), 65–91.

Goldinger, S. D. (1992). *Words and voices: Implicit and explicit memory for spoken words*. Research on speech perception (Tech. Rep. No. 7). Bloomington: Indiana University.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1166-1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251–279.

Goldinger, S. D., Pisoni, D. B., & Logan, J. S. (1991). On the nature of talker variability effects in recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 152-162.

Goto, H. **(**1971). Auditory perception by normal Japanese adults of the sounds 'l' and 'r'. *Neuropsychologia*, 9, 317–323.

Halle, M. (1985). Speculations about the representations of words in memory. In V. A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor if Peter Ladefoged* (pp. 101-114). Orlando: Academic Press.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, *24*(4), 495-522.

Hattori, K., & Iverson, P. (2009). English /r/-/l/ category assimilation by Japanese adults: individual differences and the link to identification accuracy. *Journal of the Acoustical Society of America*, *125*(1), 469-479.

Hayes-Harb, R. (2002). The perception of novel phoneme contrasts in a second language: A developmental study of native speakers of English learning Japanese singleton and geminate consonant contrasts. *Coyote Papers: Working Papers in Linguistics*, *12*, 28-41.

Hayes-Harb, R. (2005). Optimal L2 speech perception: Native speakers of English and Japanese consonant length contrasts. *Journal of Language & Linguistics, 4(1),* 1–29.

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research, 23*(1), 65-94.

Hayes-Harb, R., & Masuda, K. (2008). Development of the ability to lexically encode novel L2 phonemic contrast. *Second Language Research*, *24*(1), 5–33.

Henning, W. A. (1966). Discrimination training and self-evaluation in the teaching of pronunciation. *IRAL*, *4*(1), 7-17.

Hollich, G. H., Jusczyk, P. W., & Brent, M. (2002). Talker variability and infant word learning. Poster presented at the International Conference on Infant Studies, Toronto, Canada, April 20.

Houston, D. M. (1999). *The role of talker variability in infant word representations*. (Unpublished doctoral dissertation). Johns Hopkins University, Baltimore.

Houston, D., Jusczyk, P., & Tager, J. (1998). Talker- specificity and persistence of infants' word representations. Proceedings of the 22nd Annual Boston University Conference on Language Development (Vol. 1, pp. 385– 396). Somerville, MA: Cascadilla Press.

Houston, D., & Jusczyk, P. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1570-1582.

Houston, D. M., & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception & Performance, 29,* 1143-1153.

Ingram, J. C. L., & Park, S. G. (1998). Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners, *Journal of the Acoustical Society of America*, *103*, 1161-1174.

Iverson, P., & Evans, B. (2007). Auditory training of English vowels for first-language speakers of Spanish and German. Proc. of the 16th Intern. Congress of Phonetic Sciences, 1625-1628, Saarbrücken, Germany.

Jamieson, D. G., & Morosan, D. E. (1986). Training nonnative speech contrast in adults: Acquisition of the English/ /θ̲/-/ð̲/ contrast by francophones, *Perception and*

Psychophysics, 40, 205-215.

Jamieson, D.G., & Morosan, D. E. (1989). Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology*, *43*(1), 88-96.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego: Academic Press.

Johnson, K., & Mullennix, J. W. (1997). *Talker variability in speech processing*. San Diego: Academic Press, pp. 1–237.

Jusczyk, P., & Aslin, R. (1995). Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology, 29,* 1-23.

Jusczyk, P. W., Bertoncini, J., Bijeljac-Babic. Kennedy. L. J., & Mehler, J. (1990). The role of attention in speech perception by young infants. *Cognitive Development*, *5,* 265-286.

Jusczyk, P. W., & Derrah, C. (1987). Representation of speech sounds by young infants. *Developmental Psychology, 23,* 648-654.

Jusczyk, P. W., Kennedy, L. J., & Jusczyk. A. M. (1995). Young infants retention of information about syllables. *Infant Behavior & Development*, *IS*, 27-42.

Jusczyk, P., Pisoni, D., & Mullennix, J. (1992a). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition, 43*(3), 253-291.

Jusczyk, P., Pisoni, D., & Mullennix, J. (1992b). Effects of talker variability on speech perception by 2- month-old infants. *Cognition*, *43*(3), 253–291.

Jusczyk P. W., Jusczyk A. M., Kennedy, L. J., Schomberg, T., & Koenig, N. (1995). Young infants' retention of information about bisyllabic utterances. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 822-836.

Kenworthy, J. (1987). *Teaching English pronunciation*. Essex: Longman, pp. 73-78.

Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 1-26). New Jersey: Lawrence Erlbaum Associates.

Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, *66,* 1669-1679.

Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development, 6,* 263-285.

Ladefoged, P. (1980). What are linguistics sounds made of? *Language*, *56*, 485-502.

Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics*, *26*(2), 227-247.

Leather, J., & James, A. (1991). The acquisition of second language speech. *Studies in Second Language Acquisition*, *13*, 305-341.

Lee, J., Perrachione, T. K., Dees, T. M., & Wong, P. C. (2007). Differential effects of stimulus variability and learners' pre-existing pitch perception ability in lexical tone learning by native English speakers. In *16th Meeting of the International Congress of Phonetic Sciences*.

Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. G. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology*, *122*, 316–330.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, *94*, 1242–1255.

Lively, S. E., Pisoni, D. B., & Logan, J. S. (1992). Some effects of training Japanese listeners to identify English /r/ and /l/. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), *Speech perception, production and linguistic structure* (pp. 175-196). Tokyo/Amsterdam: Ohmsha / IOS Press.

Lively, S. E, Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/: III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, *96*(4), 2076–2087.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America, 89*, 874–886.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1993). Training listeners to perceive novel phonetic categories: How do we know what is learned? *Journal of the Acoustical Society of America*, 94(2), 1148-1151.

Logan, J. S., & Pruitt, J. S. (1995). Methodological considerations in training listeners to perceive nonnative phonemes. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 351-377). Baltimore: York.

Lucas, M. (2000). Semantic priming without association: A second look. *Journal of Verbal Learning and Verbal Behavior*, *23,* 709-733.

MacKay, I.R.A., Flege, J.E., & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics, 27*, 157-183.

MacKay, I. R. A., Meador, D., & Flege, J. E. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, *58*, 103-125.

MacKain, K., Best, C., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, *2*, 369-390.

Magnuson, J. S., & Nusbaum, H. C. (2007). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *Journal of Experimental Psychology: Human Perception and Performance*, *33,* 391–409.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognition*, *10,* 29-63.

Masuda, K., Norrix, L., & Green, K. (2001). Development of speech perception and production skills of /r/ and /l/ in Japanese learners of English. *SLAT Student Association*, *8*, 1-13.

Martin, C. S., Mullennix, J. W., Pisoni, D. B., & Sommers, W. V. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 676-684.

Matthews, J., & Brown, C. (2004). When intake exceeds input: Language specific perceptual illusions produced by L1 prosodic constraints. *International Journal of Bilingualism*, *8*(1), 5-27.

Mayo, L.H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language and Hearing Research*, *40*(3), 686.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, and Behavioral Neuroscience*, *2,* 89 -109.

McClaskey, C. L., Pisoni, D. B., & Carrell, T. D. **(**1983**)**. Transfer of training of a new linguistic contrast in voicing. *Perception and Psychophysics*, *34*, 323–330.

Mochizuki, M. (1981). The identification of /r/ and /l/ in natural and synthesized speech. *Journal of Phonetics, 9*, 283-303.

Mueller, T.H., & Niedzielski, H. (1968). The influence of discrimination training on pronunciation. *Modern Language Journal*, *52*(7), 410-416.

Moyer, A. (2007). Do language attitudes determine accent? A study of bilinguals in the USA. *Journal of Multilingual and Multicultural Development*, *28*,(6) 502-517.

Mueller, J. L. (2009). The influence of lexical familiarity on ERP responses during sentence comprehension in language learners. *Second Language Research*, *25*(1), 43-76.

Mullennix, J. W., & Pisoni, D. B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception and Psychophysics*, *47*, 379-390.

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, *85*, 365–378.

Munro, M., Flege, J., & MacKay, I. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics*, *17,* 313-334.

Newman, R. S., & Shannon, E. (2007). The effect of talker familiarity on stream segregation. *Journal of Phonetics, 35*, 85–103.

Nishi, K., & Kewley-Part, D. (2007). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research*, *50*(6), 1496-509.

Nunan, D. (1991). *Language teaching methodology*. Prentice Hall: New York.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. (Research in Speech Perception, Progress Report 10). Indiana University, Bloomington, IN: Speech Research Laboratory, Department of Psychology, pp. 357–376.

Nusbaum, H., & Morin, T. (1989). *Perceptual normalization of talker differences*. Paper presented at the 117th meeting of the Acoustical Societal of America, Syracuse, NY.

Nusbaum, H., & Morin, T. (1992). Paying attention to differences among talkers. In Y. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), *Speech perception, production, and linguistic structure* (pp. 113-134). Tokyo: IOS Press.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception and Psychophysics*, *60*(3), 355–376.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, *5*(1), 42-46.

Oldfield, R. (1966). Things, words and the brain. *Quarterly Journal of Experimental Psychology*, *18*(4), 340.

Ota, M., Hartsuiker, R. J., & Haywood, S. (2009). The KEY to the ROCK: near-homophony in nonnative visual word recognition. *Cognition*, *111*, 363- 269.

Pállier, C., Colome´, A., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, 64(3), B9-B17.

Pállier, C., Colome´, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, *12*(6), 445–449.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 309- 328.

Pater, J. (2003). The perceptual acquisition of Thai phonology by English speakers: task and stimulus effects. *Second Language Research, 19(3),* 209-223.

Pater, J., Stager, C, & Werker, J. (2004). The perceptual acquisition of phonological contrasts. *Language*, *80(3),* 384-402.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels, *Journal of the Acoustical Society of America, 24,* 175-184.

Pisoni, D. B. (1971**)**. *On the nature of categorical perception of speech sounds*. Doctoral thesis, University of Michigan.

Pisoni, D. B. (1997). Some thoughts on “normalization” in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9-32). San Diego: Academic Press.

Pisoni, D. B., Aslin, R. N., Perey, A. J., & Hennessy, B. L. **(**1982**).** Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology, 8*, 297–314.

Pisoni, D. B., & Lively, S. E. (1995). Variability and invariance in speech perception: A new look at some old problems in perceptual learning. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 433–459). Timonium, MD: York.

Pisoni, D. B., Lively, S. E., & Logan, J. S. (1994). Perceptual learning of nonnative speech contrasts: Implications for theories of speech perception. In J. C. Goodman & H. C. Nusbaum (Eds.), *Development of speech perception: The transition from recognizing speech sounds to spoken words* (pp. 121 - 166). Cambridge, MA: MIT Press.

Pruitt, J. S. (1993). Comments on "Training Japanese listeners to identify /r/ and /l/: a first report [J. S. Logan, S. E. Lively, and D.B. Pisoni, Journal of the Acoustical Society of America, *89*, 874-886(1991)". *Journal of the Acoustical Society of America*, *94*, 1146-1147.

Rochet, B. L. **(**1995**)**. Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), S*peech perception and linguistic experience: Issues in cross-language research* (pp. 379–410). Timonium, MD: York.

Rost, G., & McMurray, B. (2009) Speaker variability augments phonological processing in early word learning. *Developmental Science, 12*(2*)*, 339-349.

Rost, G., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, *15*(6), 608-635.

Ryalls, B. O., & Pisoni, D. B. (1997). The effect of talker variability on word recognition in preschool children. *Developmental Psychology*, *33*(3), May 1997, 441-452.

Salverda, A.P., Dahan, D., & McQueen, J.M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, *90*, 51–89.

Schacter, D. L., & Church, B. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 915-930.

Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: flexibility of early word representations. *Developmental Science*, *12*(4), 583-601.

Sekiguchi, T. (2006). Effects of lexical prosody and word familiarity on lexical access of spoken Japanese words. *Journal of Psycholinguistic Research, 35*, 369-384.

Shatzman, K.B., & J.M. McQueen. (2006). Prosodic knowledge affects the recognition of newly acquired words. *Psychological Science*, *17,* 5, 372-377.

Sheffert, S. M., & Fowler, C. A. (1995). The effects of voice and visible speaker change on memory for spoken words. *Journal of Memory and Language, 34*, 665-685.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics, 3*, 243-261.

Shehata, A. (2007). Exploring the production of novel Arabic phonemes by native adult English speakers. Unpublished research.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Applied Psycholinguistics*, *84*,127–190.

Shimizu, K., & Dantsuji, M. (1983). A study on the perception of *Irl* and *III* in natural and synthetic speech sounds, *Studia* Phonologica, *17*, 1-14.

Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition*, *106*, 833–870.

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: does musical ability matter? *Psychological Science, 17*(8), 675-681.

Sommers, M. S., Nygaard, L. C., & Pisoni, D. B. (1994). Stimulus variability and spoken word recognition. I. Effects of variability in speaking rate and overall amplitude. *Journal of the Acoustical Society of America, 96,* 1314-1324.

Sommers, M. S., Kirk, K. I., & Pisoni, D. B. (1997). Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners. I: The effects of response format. *Ear and Hearing*, *18*, 89-99.

Spence, M. J., Rollins, P. R., & Jerger, S. (2002). Children's recognition of cartoon voices. *Journal of Speech, Language, and Hearing Research, 45*(1), 214-222.

SPSS, Inc. (2005). *SPSS* (Version 13.0). Chicago, IL: John Wiley & Sons.

Stockmal, V., Moates, D. R., & Bond, Z. S. (2000). Same talker, different language. *Applied Psycholinguistics*, *21*(3), 383-393.

Strange, W. (1972). *The effects of training on the perception of synthetic speech sounds: Voice onset time.* Unpublished doctoral dissertation, University of Minnesota.

Strange, W., & Dittman, S. (1984). Effect of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception and Psychophysics*, *36*(2), 131-145.

Strange, W., & Jenkins, J. (1978). Role of linguistic experience in perception of speech. In R. D. Walk & H. L. Pick (Eds.), *Perception and experience* (pp. 125–169). New York: Plenum Press.

Strange, W., Polka, L., & Aguilar, M. C. (1989). Effects of auditory and phonetic training on American's discrimination of Hindi retroflex- dental contrasts. *The Journal of the Acoustical Society of America Supplement*, *I*, 86, SI01.

Strange, W. (1995). Phonetics of second language acquisition: Past, present, future. *ICPhS*, *95*(*4)*, 76-83.

Streeter, L. A., & Landauer, T. K. (1976). Effects of learning English as a second language on the acquisition of a new phonemic contrast. *Journal of the Acoustical Society of America*, *59*, 448-451.

Summerfield, Q., & Haggard, M. P. (1973). Report on Research in Progress in Speech Perception No. 2. Belfast, Northern Ireland: Queen's University of Belfast. Vocal tract normalization as demonstrated by reaction times.

Sumner, M. (2011). The role of variation in the perception of accented speech. *Cognition*, *119*, 131-136.

Tajima, K., Rothwell, A., & Munhall, K.G. (2002). Native and nonnative perception of phonemic length contrasts in Japanese: Effect of identification training and exposure*. Journal of the Acoustical Society of America*, *112*, 2387.

Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America, 111*(*6)*, 2887-2896.

Takagi, N., & Mann, V. **(**1995). The limits of extended naturalistic exposure on the perceptual mastery of English /r/ and /l/ by adult Japanese learners of English. *Applied Psycholinguistics*, *16*, 379–40.

Tincoff, R., & Jusczyk, P. W. (1996). *Are word-final sounds perceptually salient for infants?* Poster presented at the Fifth Conference on Laboratory Phonology, Evanston, FL.

Theodore, R. M., & Miller, J. L. (2010). Characteristics of listener sensitivity to talker-specific phonetic detail. *Journal of the Acoustical Society of America, 128*(4), 2090-2099.

Vitevitch, M. S., & Rodríguez, E. (2005). Neighbourhood density effects in spoken word recognition in Spanish. *Journal of Multilingual Communication Disorders*, *3*, 64–73.

Wade, T., Jongman, A., & Sereno, J. (2007). Effects of acoustic variability in the perceptual learning of nonnative-accented speech sounds. *Phonetica, 64,*122-144.

Wang, Y., Jongman, A., & Sereno, J. A . (2006). Second language acquisition and processing of Mandarin tone. In E. Bates, L. Tan, & O. Tzeng (Eds.), *Handbook of Chinese psycholinguistics* (pp. 250-257). Cambridge: Cambridge University Press.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America*, *113*(2), 1033-1043.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America*, *106*(6), 3649-3658.

Weber, A., & Cutler, A. (2004). Lexical competition in nonnative spoken-word recognition. *Journal of Memory and Language*, 50, 1–25.

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Development aspects of cross-language speech perception. *Child Development*, *52*(1), 349-355.

Werker, J. F., & Tees, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7,*49-63.

Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*(6), 1866-1878.

Wong, P. C. M., & Perrachione, T. K. 2007. Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*(4), 565-585.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422.

Yamada, R. A. (1993). Effect of extended training on /r/ and /l/ identification by native speakers of Japanese. *Journal of the Acoustical Society of America*, *93,* 2391.

Yamada, R. A., Strange, W. Magnuson, J. S., Pruitt, J. S., & Clarke III, W.D. (1994). The intelligibility of Japanese speakers' productions of American English /r/, /l/, and /w/, as evaluated by native speakers of American English. *Proceedings of the International Conference of Spoken Language Processing* (Acoustical Society of Japan, Yokohama). 2023-2026.

Yamada, R., Tohkura, Y., & Kobayashi, N. (1992). Effect of word familiarity on nonnative phoneme perception: Identification of English /r/, /l/, and /w/ by native speakers of Japanese. In A. James & J. Leather (Eds.), *Second language speech* (pp. 103-117). Mouton. de Gruyter: The Hague.

Yonan, C. A., & Sommers, M. S. (2000). The effects of talker familiarity on spoken word identification in younger and older listeners. *Psychology and Aging, 15,* 88-99.

Yu, K., & Jamieson, D. G. (1993). Training of the English /r/ and /l/ speech contrasts in Korean listeners. *Canadian Acoustics*, *21*, 107-108.

Zaba, A., Bolewicz, K., & Hayes-Harb, R. (n.d.). Native English speakers perceiving Arabic pharyngealization contrasts: The effects of vowel context. Unpublished manuscript.

Zubick, H. H., Irizarry, L. M., Rosen, L., Feudo, P. Jr., Kelly, J. H., & Strome, M. (1983). Development of speech-audiometric materials for native Spanish- speaking adults. *Audiology, 22,* 88-102.